OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 17, 2004, 04:56:59 ; Search time 2791.11 Seconds
                                        (without alignments)
                                        10714.982 Million cell updates/sec

Title: US-09-367-052-3
Perfect score: 690
Sequence: 1 ctgcacctgtcagtggctga........gcctgaacccccatcctctat 690

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
          1:   gb_ba:*
          2:   gb_htg:*
          3:   gb_in:*
          4:   gb_om:*
          5:   gb_ov:*
          6:   gb_pat:*
          7:   gb_ph:*
          8:   gb_pl:*
          9:   gb_pr:*
         10:   gb_ro:*
         11:   gb_sts:*
         12:   gb_sy:*
         13:   gb_un:*
         14:   gb_vi:*
         15:   em_ba:*
         16:   em_fun:*
         17:   em_hum:*
         18:   em_in:*
         19:   em_mu:*
         20:   em_om:*
         21:   em_or:*
         22:   em_ov:*
         23:   em_pat:*
         24:   em_ph:*
         25:   em_pl:*
         26:   em_ro:*
         27:   em_sts:*
         28:   em_un:*
         29:   em_vi:*
         30:   em_htg_hum:*
         31:   em_htg_inv:*
         32:   em_htg_other:*
         33:   em_htg_mus:*
         34:   em_htg_pln:*
         35:   em_htg_rod:*
         36:   em_htg_mam:*
         37:   em_htg_vrt:*
         38:   em_sy:*
         39:   em_htgo_hum:*
         40:   em_htgo_mus:*
         41:   em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	690	100.0	1180	10	MMLCR12	Z80112 Mus musculu
2	690	100.0	1817	10	BC031665	BC031665 Mus muscu
3	690	100.0	1877	10	D87747	D87747 Mus musculu
4	688.4	99.8	1422	10	MMLCR13	Z80113 Mus musculu
5	685.2	99.3	1223	10	MMU59760	U59760 Mus musculu
6	685.2	99.3	3366	10	MMU65580	U65580 Mus musculus
7	682	98.8	1575	10	MMLESTRPT	X99582 M.musculus
8	682	98.8	1809	10	AB000803	AB000803 Mouse mRN
9	682	98.8	3770	10	MMLESTRGN	X99581 M.musculus
10	682	98.8	175370	2	AC145861	AC145861 Pan trogl
11	682	98.8	185215	2	AC147556	AC147556 Mus muscu
12	578.6	83.9	1050	10	AF452185	AF452185 Rattus no
13	578.6	83.9	190149	2	AC122097	AC122097 Rattus no
14 c	578.6	83.9	240931	2	AC107097	AC107097 Rattus no
15	575.4	83.4	1050	10	RNU90610	U90610 Rattus norv
16	523.6	75.9	1059	9	AF031089	AF031089 Papio ham
17	520.4	75.4	1044	9	AF172212	AF172212 Pygathrix
18	518.8	75.2	1003	9	AF178084	AF178084 Saguinus
19	518.8	75.2	1044	9	AF172210	AF172210 Pygathrix
20	518.8	75.2	1044	9	AF172211	AF172211 Rhinopith
21	518.8	75.2	1044	9	AF172213	AF172213 Pygathrix
22	518.8	75.2	1044	9	AF172214	AF172214 Nasalis l
23	518.8	75.2	1044	9	AF172218	AF172218 Presbytis
24	518.8	75.2	1044	9	AF172227	AF172227 Hylobates
25	518.8	75.2	1044	9	AF172228	AF172228 Hylobates
26	518.8	75.2	1044	9	AF172231	AF172231 Pongo pyg
27	518.8	75.2	1059	9	AF452613	AF452613 Saimiri s
28	517.2	75.0	901	9	AY185894	AY185894 Macaca fa
29	517.2	75.0	1003	9	AF229128	AF229128 Saimiri b
30	517.2	75.0	1044	9	AF178082	AF178082 Cercopith
31	517.2	75.0	1044	9	AF172219	AF172219 Presbytis
32	517.2	75.0	1044	9	AF172220	AF172220 Presbytis
33	517.2	75.0	1044	9	AF172221	AF172221 Mandrillu
34	517.2	75.0	1044	9	AF172223	AF172223 Macaca as
35	517.2	75.0	1044	9	AF172224	AF172224 Macaca ar
36	517.2	75.0	1044	9	AF172225	AF172225 Macaca th
37	517.2	75.0	1044	9	AF172226	AF172226 Macaca ne
38	517.2	75.0	1044	9	AF172229	AF172229 Hylobates
39	517.2	75.0	1044	9	AF172230	AF172230 Hylobates
40	517.2	75.0	1059	6	AX256185	AX256185 Sequence
41	517.2	75.0	1059	6	AX482330	AX482330 Sequence
42	517.2	75.0	1059	6	BD226095	BD226095 Therapeut
43	517.2	75.0	1059	9	AF025375	AF025375 Homo sapi
44	517.2	75.0	1059	9	AY242129	AY242129 Homo sapi
45	517.2	75.0	1059	9	BT006660	BT006660 Homo sapi

ALIGNMENTS

RESULT 1
MMLCR12
LOCUS   MMLCR12   1180 bp   mRNA   linear   ROD 18-NOV-1997
DEFINITION Mus musculus lcr-1 gene.
ACCESSION   Z80112
VERSION   Z80112.1 GI:2632100
KEYWORDS   CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1180)
AUTHORS   Schubel,A., Burgstahler,R. and Lipp,M.
TITLE   The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
            High expression in thymus and lymphoid tissues

Right column:

Db	592 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGTCTCATCCTGCCCGGC 651
Qy	421 ATGTGCTCATCCTCCCTGCTTACTGGTCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 480
Db	652 ATCGTCATCCTCCCTGCTTACTGGTCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 711
Qy	481 CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCATCCTCAGCTTTCTTGCCTGCTGG 540
Db	712 CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCATCCTCAGCTTTCTTGCCTGCTGG 771
Qy	541 CTGCCATATTATGTGGGATCAGCATCGACTCCTTCCTTTGGAGTCATCAAGCAA 600
Db	772 CTGCCATATTATGTGGGATCAGCATCGACTCCTTCCTTTGGAGTCATCAAGCAA 831
Qy	601 GGATGGACTTCGAGAGCATTGTGCACAAGTGGATCCATCACAGAGGCCCTCGCCTTC 660
Db	832 GGATGGACTTCGAGAGCATTGTGCACAAGTGGATCCCATCACAGAGGCCCTCGCCTTC 891
Qy	661 TTCCACTGTTGCCTGACCCCATCCTCTAT 690
Db	892 TTCCACTGTTGCCTGACCCCATCCTCTAT 921

```
RESULT 2
BC031665
LOCUS       BC031665    1817 bp    mRNA    linear    ROD 07-OCT-2003
DEFINITION  Mus musculus chemokine (C-X-C motif) receptor 4, mRNA (cDNA clone
            MGC:36266 IMAGE:3592479), complete cds.
ACCESSION   BC031665
VERSION     BC031665.1  GI:21618781
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1817)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
  PUBMED    12477932
REFERENCE   2  (bases 1 to 1817)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: cgapbs-r@mail.nih.gov
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (iLLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
```

Left column:

```
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 1180)
  AUTHORS   Lipp,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
            Molecular Medicine, MDC, Robert-Roessle-strasse 10, BERLIN-BUCH,
            D-13122, GERMANY
COMMENT     On Nov 21, 1997 this sequence version replaced gi:1542888.
FEATURES             Location/Qualifiers
     source          1..1180
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129SV"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
                     /dev_stage="adult"
     gene            1..1180
                     /gene="lcr-1"
     CDS             1..1074
                     /gene="lcr-1"
                     /function="receptor for chemokine SDF-1"
                     /standard_name="CXC chemokine receptor 4"
                     /codon_start=1
                     /product="CXCR-4"
                     /protein_id="CAB02202.1"
                     /db_xref="GI:1542889"
                     /db_xref="GOA:P70658"
                     /db_xref="SWISS-PROT:P70658"
                     /translation="MEPISIYTSDNYSEVGSGGDYDSNKEPCFRDENVHFNRIFLPTI
                     YFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADW
                     YFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVVGV
                     WIPALLLTIPDFIFADVSGGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGI
                     VILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGVIK
                     QGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKIL
                     SKGKRGGHSSVSTESESSSFHSS"
     exon            <1..15
                     /gene="lcr-1"
                     /number=1
     exon            16..>1180
                     /gene="lcr-1"
                     /number=2
ORIGIN

Query Match       100.0%;  Score 690;  DB 10;  Length 1180;
Best Local Similarity  100.0%;  Pred. No. 7.2e-163;
Matches 690;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

Qy	1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
Db	232 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 291
Qy	61 GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT 120
Db	292 GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT 351
Qy	121 GTCAACCTCTACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
Db	352 GTCAACCTCTACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 411
Qy	181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
Db	412 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 471
Qy	241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
Db	472 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 531
Qy	301 AGCGGAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTACCCCGAT 360
Db	532 AGCGGAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTACCCCGAT 591
Qy	361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGTCTCATCCTGCCCGGC 420

```
Qy  301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCGTGACCGGCCTTTACCCGAT 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  615 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCGTGACCGGCCTTTACCCGAT 674

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  675 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC 734

Qy  421 ATCGTCATCCTCTCCTGTTACTGACATCATCATCCTAAGCGTCACACTCCAAGGGCCAC 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  735 ATCGTCATCCTCTCCTGTTACTGACATCATCATCCTAAGCGTCACACTCCAAGGGCCAC 794

Qy  481 CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTCATCCTCTTTGCCTGCTGG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  795 CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTCATCCTCTTTGCCTGCTGG 854

Qy  541 CTGCCATATTATGTGGGATCAGCATGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  855 CTGCCATATTATGTGGGATCAGCATGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 914

Qy  601 GGGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTGCCTTC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  915 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTGCCTTC 974

Qy  661 TTCCACTGTTGCCTGAACCCCATCCTCAT 690
        |||||||||||||||||||||||||||||
Db  975 TTCCACTGTTGCCTGAACCCCATCCTCAT 1004
```

RESULT 3
D87747
```
LOCUS       D87747                  1877 bp    mRNA     linear   ROD 07-FEB-1999
DEFINITION  Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION   D87747
VERSION     D87747.1  GI:1772445
KEYWORDS    murine CXCR-4.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (sites)
  AUTHORS   Nagasawa,T., Nakajima,T., Tachibana,K., Iizasa,H., Bleul,C.C.,
            Yoshie,O., Matsushima,K., Yoshida,N., Springer,T.A. and
            Kishimoto,T.
  TITLE     Molecular cloning and characterization of a murine pre-B-cell
            growth-stimulating factor/stromal cell-derived factor-1 receptor, a
            murine homolog of the human immunodeficiency virus 1 entry
            coreceptor fusin
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)
   MEDLINE  97121456
   PUBMED   8962122
REFERENCE   2  (bases 1 to 1877)
  AUTHORS   Nagasawa,T.T.
  TITLE     Molecular cloning of murine PBSF/SDF-1 receptor
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1877)
  AUTHORS   Nagasawa,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-SEP-1996) Takashi Nagasawa, Research Institute, Osaka
            Medical Center, Department of Immunology; Murodocho 840, Izumi,
            Osaka 590-02, Japan (E-mail:immunol@osk.threewebnet.or.jp,
            Tel:0725-56-1220, Fax:0725-57-3021)
FEATURES             Location/Qualifiers
     source          1..1877
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="DW34"
                     /cell_type="pre-B cell"
                     /tissue_type="bone marrow"
     gene            1..1877
                     /gene="murine CXCR-4"
```

```
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 61 Row: d Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES             Location/Qualifiers
     source          1..1817
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:36266 IMAGE:3592479"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Man1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1..1817
                     /gene="Cxcr4"
                     /note="synonyms: LESTR, fusin, PB-CKR, PBSF/SDF-1, CD184"
                     /db_xref="LocusID:12767"
                     /db_xref="MGI:109563"
     CDS             78..1157
                     /codon_start=1
                     /product="chemokine (C-X-C motif) receptor 4"
                     /protein_id="AAH31665.1"
                     /db_xref="GI:21618782"
                     /db_xref="LocusID:12767"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIFLITGIVGNGLVILVMGVYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFBSIVHKWISITEALAFFHCCLNPIIYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSFHSS"
     misc_feature    246..1004
                     /note="7tm_1; Region: 7 transmembrane receptor (rhodopsin
                     family)"
                     /db_xref="CDD:pfam00001"
ORIGIN
Query Match            100.0%;  Score 690;  DB 10;  Length 1817;
Best Local Similarity  100.0%;  Pred. No. 7.3e-163;
Matches 690;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy    1 CTGCCACCTGCTCAGTGGCTGACCTGCCCCTTGTCATCACACCCCCTTCTGGGCAGTTGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  315 CTGCCACCTGCTCAGTGGCTGACCTGCCCCTTGTCATCACACCCCCTTCTGGGCAGTTGAT 374

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  375 GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT 434

Qy  121 GTCAACCTCACAGACAGCAGCGGTTCTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  435 GTCAACCTCACAGACAGCAGCGGTTCTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGCC 494

Qy  181 ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAGGCAGTCTAT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  495 ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAGGCAGTCTAT 554

Qy  241 GTGGGGCGTCTGGATCCCAGCCCTCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  555 GTGGGGCGTCTGGATCCCAGCCCTCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 614
```

dB-09-367/-052-3.rge

```
CDS              120..1199
                 /gene="murine CXCR-4"
                 /function="PBSF/SDF-1 receptor"
                 /codon_start=1
                 /evidence=experimental
                 /product="murine CXCR-4"
                 /protein_id="BAA13451.1"
                 /db_xref="GI:1772446"
                 /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                 TIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                 DWYFGKFLCKAVHIIYTVNLYSSVLLIAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                 GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                 GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
                 IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                 ILSKGKRGGHSSVSTESESSSFHSS"

ORIGIN

Query Match        100.0%;   Score 690;  DB 10;  Length 1877;
Best Local Similarity  100.0%;   Pred. No. 7.3e-163;
Matches  690;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy       1 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCGGGCAGTTGAT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     357 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCGGGCAGTTGAT  416

Qy      61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     417 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  476

Qy     121 GTCAACCTCTACAGACAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCTGGTACCTCGCC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     477 GTCAACCTCTACAGACAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCTGGTACCTCGCC  536

Qy     181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     537 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  596

Qy     241 GTGGGGGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     597 GTGGGGGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  656

Qy     301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGACCGGCCTTACCCCGAT  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     657 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGACCGGCCTTACCCCGAT  716

Qy     361 AGCCTGTGAGATGGTGGTGTTTCAATTCCAGCATATAATGTGGGTCTCATCCTGCCCGGC  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     717 AGCCTGTGAGATGGTGGTGTTTCAATTCCAGCATATAATGTGGGTCTCATCCTGCCCGGC  776

Qy     421 ATGGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     777 ATGGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC  836

Qy     481 CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTGGCTTTCTTTGCCTGCTGG  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     837 CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTGGCTTTCTTTGCCTGCTGG  896

Qy     541 CTGCCATATATGTGGGAGATCAGCATCAAGACTCCTCCATCCTTTGGGAGTCATCAAGCAA  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     897 CTGCCATATATGTGGGAGATCAGCATCAAGACTCCTCCATCCTTTGGGAGTCATCAAGCAA  956

Qy     601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGCCCTTCGCCTTC  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     957 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGCCCTTCGCCTTC  1016

Qy     661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
           ||||||||||||||||||||||||||||||
Db    1017 TTCCACTGTTGCCTGAACCCCATCCTCTAT  1046

RESULT 4
MMLCR13
LOCUS    MMLCR13     1422 bp    DNA    linear    ROD 31-OCT-1996
```

```
DEFINITION  Mus musculus lcr-1 gene, exon 2.
ACCESSION   Z80113
VERSION     Z80113.1  GI:1655638
KEYWORDS    CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1422)
  AUTHORS   Schubel,A., Burgstahler,R. and Lipp,M.
  TITLE     The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
            High expression in thymus and lymphoid tissues
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1422)
  AUTHORS   Lipp,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
            Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
            D-13122, GERMANY
FEATURES             Location/Qualifiers
     source          1..1422
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV"
                     /db_xref="taxon:10090"
                     /clone_lib="lambda fix (TM) II"
                     /dev_stage="adult"
     intron          <1..259
                     /number=1
     gene            260..1422
                     /gene="lcr-1"
     exon            260..>1422
                     /gene="lcr-1"
                     /product="CXCR-4"
                     /standard_name="CXC chemokine receptor 4"
                     /number=2
                     /function="receptor for chemokine SDF-1"

ORIGIN

Query Match        99.8%;    Score 688.4;  DB 10;  Length 1422;
Best Local Similarity  99.9%;    Pred. No. 1.8e-162;
Matches  689;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy       1 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     476 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT  535

Qy      61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     536 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  595

Qy     121 GTCAACCTCTACAGACAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCTGGTACCTCGCC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     596 GTCAACCTCTACAGACAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCTGGTACCTCGCC  655

Qy     181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     656 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  715

Qy     241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     716 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  775

Qy     301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGGCCTTACCCCGAT  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     776 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGGCCTTACCCCGAT  835

Qy     361 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     836 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC  895

Qy     421 ATGGTCATCCTCCTGTTACTGCTACATCATCATCTCCTAAGCTGTCACACTCCAAGGGCCAC  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   896 ATCGTCATCCTCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 955

Qy   481 CAGAAGCGCAAGGCCCTCAAGACGACGAGTCATCCTCATCCTCCTCCAGCTTCTTTGCCTGCTGG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   956 CAGAAGCGCAAGGCCCTCAAGACGACGAGTCATCCTCATCCTCCTCCAGCTTCTTTGCCTGCTGG 1015

Qy   541 CTGCCATTATGTGGGGATCAGCAGCATCGACTCCTCCTTCATCCTTTGGGAGTGCATCAAGCAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1016 CTGCCATTATGTGGGGATCAGCAGCATCGACTCCTCCTTCATCCTTTGGGAGTGCATCAAGCAA 1075

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAAGTGGATCTCCGCCTTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1076 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAAGTGGATCTCCGCCTTC 1135

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
         ||||||||||||||||||||||||||||||
Db 1136 TTCCACTGTTGCCTGAACCCCATCCTCTAT 1165

RESULT 5
MMU59760
LOCUS   MMU59760   1223 bp   mRNA   linear   ROD 30-SEP-2002
DEFINITION   Mus musculus fusin homolog mRNA, complete cds.
ACCESSION   U59760
VERSION   U59760.1 GI:1527134
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1223)
  AUTHORS   Heesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
  TITLE   Cloning of the mouse fusin gene, homologue to a human HIV-1
             co-factor
  JOURNAL   J. Immunol. 157 (12), 5455-5460 (1996)
   MEDLINE   97113334
   PUBMED   8955194
REFERENCE   2 (bases 1 to 1223)
  AUTHORS   Heesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1996) Pathology, Harvard Medical School, 200
             Longwood Avenue, Boston, MA 02115, USA
FEATURES   Location/Qualifiers
     source   1..1223
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129/Sv"
                     /db_xref="taxon:10090"
     CDS   30..1109
                     /note="seven transmembrane spanning receptor;
                     chemokine-receptor like"
                     /codon_start=1
                     /product="fusin homolog"
                     /protein_id="AAB07725.1"
                     /db_xref="GI:1527135"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWTFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILLLAFFACWLPYYVGISIDSFILLGV
                     IKQCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHS"
ORIGIN
Query Match   99.3%; Score 685.2; DB 10; Length 1223;
Best Local Similarity   99.6%; Pred. No. 1.2e-161;
Matches 687; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   267 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 326

Qy   61 GCCATGGCTGACTGGACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120

Db   327 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 386

Qy   121 GTCAACCTCTACAGACGTTCTCATCCTGCCTTCATCAGCCTGGACCCGGTACCTCGCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   387 GTCAACCTCTACAGACGTTCTCATCCTGCCTTCATCAGCCTGGACCCGGTACCTCGCC 446

Qy   181 ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGTGAAAAGGCAGTCTAT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   447 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGTGGCTGAAAAGGCAGTCTAT 506

Qy   241 GTGGGGTCTGGGATCCCAGCCCTCCTGCTGACTATACCTGACATCTTTGCCGACGTC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   507 GTGGGGTCTGGGATCCCAGCCCTCCTGCTGACTATACCTGACATCTTTGCCGACGTC 566

Qy   301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGTACATCGTGACCGGTCTCATCCTGCCCGGC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 AGCCAGGGGGACATCAGTCAGGGGGATGACAGTACATCGTGACCGGTCTCATCCTGCCCGGC 626

Qy   361 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   627 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCTCACACTCCAAGGGCCAC 686

Qy   421 ATCGTCATCCTCCTCGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   687 ATCGTCATCCTCCTCGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 746

Qy   481 CAGAAGCGCAAGGCCCTCAAGACGACGAGTCATCATCCTCATCCTCCTCCAGCTTCTTTGCCTGCTGG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   747 CAGAAGCGCAAGGCCCTCAAGACGACGAGTCATCATCCTCATCCTCCTCCAGCTTCTTTGCCTGCTGG 806

Qy   541 CTGCCATTATGTGGGGATCAGCAGCATCGACTCCTCCTTTGGGAGTCATCAAGCAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   807 CTGCCATTATGTGGGGATCAGCAGCATCGACTCCTCCTTTGGGAGTCATCAAGCAA 866

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGGCCCTCGCCTTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   867 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGGCCCTCGCCTTC 926

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
         ||||||||||||||||||||||||||||||
Db   927 TTCCACTGTTGCCTGAACCCCATCCTCTAT 956

RESULT 6
MMU65580
LOCUS   MMU65580   3366 bp   DNA   linear   ROD 13-DEC-1996
DEFINITION   Mus musculus fusin (CXCR-4) gene, complete cds.
ACCESSION   U65580
VERSION   U65580.1 GI:1731650
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 3366)
  AUTHORS   Heesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
  TITLE   Cloning of the mouse fusin gene, homologue to a human HIV-1
             co-factor
  JOURNAL   J. Immunol. 157 (12), 5455-5460 (1996)
   MEDLINE   97113334
   PUBMED   8955194
REFERENCE   2 (bases 1 to 3366)
  AUTHORS   Heesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
  TITLE   Direct Submission
  JOURNAL   Submitted (30-JUL-1996) Pathology, Harvard Medical School, 200
             Longwood Avenue, Boston, MA 02115, USA
FEATURES   Location/Qualifiers
     source   1..3366
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"

```
Qy   541  CTGCCATATTATGTGGGATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3064  CTGCCATATTATGTGGGATCAGCATCAGCATCGACTCCTTCATCCTTTGGGGGTCATCAAGCAA  3123

Qy   601  GGATGTGACTTCGAGAGACATTGTGCACAAGTGATCTCCATCACAGAGCCCTCGCCTTC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3124  GGATGTGACTTCGAGAGACATCGTGCACAAGTGGATCTCCATCACAGAGCCCTCGCCTTC  3183

Qy   661  TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
          ||||||||||||||||||||||||||||||
Db  3184  TTCCACTGTTGCCTGAACCCCATCCTCTAT  3213
```

RESULT 7

LOCUS   MMLESTRPT   1575 bp   mRNA   linear   ROD 11-NOV-1996
DEFINITION M.musculus mRNA for leukocyte-derived seven transmembrane domain
            receptor.
ACCESSION   X99582
VERSION   X99582.1 GI:1666648
KEYWORDS   lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Moepps,B., Frodl,R., Kessler,H. and Gierschik,P.
  TITLE   cDNA cloning and genomic organization of a leukocyte-derived seven
            transmembrane domain receptor (LESTR) from mouse: a murine
            homologue of the human HIV-1 entry cofactor fusin
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1575)
  AUTHORS   Moepps,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JUL-1996) B. Moepps, Universitet Ulm,
            Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK   Revised by author 11-NOV-96
COMMENT   On Nov 12, 1996 this sequence version replaced gi:1657351.
FEATURES   Location/Qualifiers
     source   1..1575
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="B6CBA"
                     /db_xref="taxon:10090"
                     /haplotype="dip_loid"
                     /sex="female"
                     /tissue_type="thymus"
                     /dev_stage="6-8 weeks old"
     gene   1..1575
                     /gene="lestr"
     CDS   90..1169
                     /gene="lestr"
                     /codon_start=1
                     /product="leukocyte-derived seven transmembrane domain
                     receptor"
                     /protein_id="CAA67894.1"
                     /db_xref="GI:1666649"
                     /db_xref="GOA:P70658"
                     /db_xref="SWISS-PROT:P70658"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLPCKAVHIIYTVNLYSSVLLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLVLP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"
     polyA_signal   1498..1503
                     /gene="lestr"

ORIGIN

Query Match   98.8%;   Score 682; DB 10; Length 1575;
Best Local Similarity 99.3%;   Pred. No. 7.5e-161;
Matches 685;   Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

Left column:

```
gene            1..3366
                /gene="CXCR-4"
CDS             join(1..21,2308..3366)
                /gene="CXCR-4"
                /note="SDF-1 chemokine receptor"
                /codon_start=1
                /product="fusin"
                /protein_id="AAC52953.1"
                /db_xref="GI:3645915"
                /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                TIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                DWYFGKFLPCKAVHIIYTVNLYSSVLLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISDFSFILLGV
                IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                ILSKGKRGGHSSVSTESESSSFHSS"
exon            1..21
                /gene="CXCR-4"
                /number=1
exon            1..15
                /gene="CXCR-4"
                /note="alternate exon 1"
intron          16..>2307
                /gene="CXCR-4"
                /note="alternate intron 1"
intron          22..2307
                /gene="CXCR-4"
                /number=1
exon            2308..3366
                /gene="CXCR-4"
                /number=2

ORIGIN

Query Match         99.3%;    Score 685.3;  DB 10;  Length 3366;
Best Local Similarity  99.6%;    Pred. No. 1.2e-161;
Matches 687;   Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;
```

```
Qy     1  CTGCACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTCGGCAGTTGAT  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2524  CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTCGGCAGTTGAT  2583

Qy    61  GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2584  GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT  2643

Qy   121  GTCAACCTCTACAGCACGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2644  GTCAACCTCTACAGCACGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  2703

Qy   181  ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2704  ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  2763

Qy   241  GTGGGGCGTCTGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2764  GTGGGGCGTCTGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  2823

Qy   301  AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCCTTACCCCGAT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2824  AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT  2883

Qy   361  AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATAATAATGTGGGTCTCATCCTGCCCGGC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2884  AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATAATAATGGTGGGTCTCATCCTGCCCGGC  2943

Qy   421  ATCGTGCATCCTCCTGTTACTGCATCATCATCCTCAAGCTGTCCACACTCCAAGGGCCAC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2944  ATCGTCATCCTCCTCTGTTACTGCATCATCATCCTCCAAGCTGTCCACACTCCAAGGGCCAC  3003

Qy   481  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTGGCTTTCTTTGCCTGCTGG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3004  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGG  3063
```

FEATURES   Location/Qualifiers
source   Fax:043-284-1736)
                     1..1809
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL-6"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
                     /dev_stage="E17"
CDS   72..1151
                     /note="fusin"
                     /codon_start=1
                     /product="murine CXCR-4"
                     /protein_id="BAA19187.1"
                     /db_xref="GI:1816446"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHNRIFLP
                     TIYFIIFLTGIVGNGLVLVLMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLLLAFISLDRYLAIVHATNSQPRRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLAMVVFQFQHIMVGLVLP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVIILAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQAHLNSMRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"

ORIGIN

Query Match   98.8%; Score 682; DB 10; Length 1809;
Best Local Similarity   99.3%; Pred. No. 7.6e-161;
Matches   685;   Conservative   0;   Mismatches   0;   Indels   5;   Gaps   0;

```
Qy      1 CTGCACCTGTCAGTGGCTGACCCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    309 CTGCACCTGTCAGTGGCTGACCCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 368

Qy     61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    369 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 428

Qy    121 GTCAACCTCTACAGCAGCAGCGTTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    429 GTCAACCTCTACAGCAGCAGCGTTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGCT 488

Qy    181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    489 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 548

Qy    241 GTGGGGCGTCTGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    549 GTGGGGCGTCTGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 608

Qy    301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    609 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 668

Qy    361 AGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCCGGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    669 AGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCTCGTCCTGCCCCGGC 728

Qy    421 ATCGTCATCCTCCTGTTACTGCATCATCATCCTAAGCTGTCCATCATCATCTACACTGTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    729 ATCGTCATCCTCCTGTTACTGCATCATCATCCTAAGCTGTCCATCATCATCTACACTGTC 788

Qy    481 CAGAAGGCAAGGCCCTCAAGACGACAGCATCGACCATCCTCATCCTTCTTGCCTGCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    789 CAGAAGCGCAAGGCCCTCAAGACGACAGCATCATCCTCATCCTTCTTTGCCTGCTGG 848

Qy    541 CTGCCATATTATGTGGGGTCATCAAGCAATCAGCTTTGGGAGTCATCAAGCAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    849 CTGCCATATTATGTGGGGTCATCAAGCATCGACTCCTTTGGGGGTCATCAAGCAA 908

Qy    601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTTGCCTTC 660
          
Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
```

RESULT 8
AB000803
LOCUS   AB000803   1809 bp   mRNA   linear   ROD 05-FEB-1999
DEFINITION Mouse mRNA for murine CXCR-4, complete cds.
ACCESSION   AB000803
VERSION   AB000803.1 GI:1816445
KEYWORDS   murine CXCR-4; fusin.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1809)
AUTHORS   Suzuki,G., Nakata,Y., Uzawa,A., Shirasawa,T., Saito,T. and Mita,K.
TITLE   Molecular cloning of murine CXCR-4, a murine homologue of HIV entry
            co-receptor, that is expressed on leukocyte and other organs
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1809)
AUTHORS   Suzuki,G.
TITLE   Direct Submission
JOURNAL   Submitted (29-JAN-1997) Gen Suzuki, National Institute of
            Radiological Sciences, Division of Radiation Health; 4-9-1,
            Anagawa, Inage-ku, Chiba, Chiba 263, Japan
            (E-mail:gsuzuki@unmxl.nirs.go.jp, Tel:043-251-2111,

```
Qy      1 CTGCACCTGTCAGTGGCTGACCCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 CTGCACCTGTCAGTGGCTGACCCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 386

Qy     61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 446

Qy    121 GTCAACCTCTACAGCAGCAGCGTTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 GTCAACCTCTACAGCAGCAGCGTTCTTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGCT 506

Qy    181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 ATTGTCCACGCCACCAACAGTCAGAGAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 566

Qy    241 GTGGGGCGTCTGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    567 GTGGGGCGTCTGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 626

Qy    301 AGCCAGGGGGACACTCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    627 AGCCAGGGGGACACTCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 685

Qy    361 AGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCTCGTCCTGCCCCGGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCTCGTCCTGCCCCGGC 746

Qy    421 ATCGTCATCCTCCTGTTACTGCATCATCATCCTAAGCTGTCCATCATCATCTACACTGTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    747 ATCGTCATCCTCCTGTTACTGCATCATCATCCTAAGCTGTCCATCATCATCTACACTGTC 806

Qy    481 CAGAAGGCAAGGCCCTCAAGACGACAGCATCGACCATCCTCATCCTTCTTTGCCTGCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    807 CAGAAGCGCAAGGCCCTCAAGACGACAGCATCATCCTCATCCTTCTTTGCCTGCTGG 866

Qy    541 CTGCCATATTATGTGGGGTCATCAAGCAATCAGCTTTGGGAGTCATCAAGCAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    867 CTGCCATATTATGTGGGGTCATCAAGCATCGACTCCTTTGGGGGTCATCAAGCAA 925

Qy    601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTTGCCTTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    927 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTTGCCTTC 986

Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
          ||||||||||||||||||||||||||||||
Db    987 TTCCACTGTTGCCTGAACCCCATCCTCTAT 1016
```

Query Match   98.8%;   Score 682;   DB 10;   Length 3770;
Best Local Similarity   99.3%;   Pred. No. 7.8e-161;
Matches 685;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy     1 CTGCACCTGTCAGTGCTGACCTCCTTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2522 CTGCACCTGTCAGTGCTGACCTCCTTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 2581

Qy    61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2582 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 2641

Qy   121 GTCAACCTCTACAGACAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2642 GTCAACCTCTACAGACAGCGGTTCTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGCT 2701

Qy   181 ATTGTCCACGCCACCACAGTCAAAGGCCGTGGCTGAAAAGGCAGTCTAT  240
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2702 ATTGTCCACGCCACCAACAGTCAGAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 2761

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTCGACTATACCTGACTTCATCTTTGCCGACGTC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2762 GTGGGCGTCTGGATCCCAGCCCTCCTCGACTATACCTGACTTCATCTTTGCCGACGTC 2821

Qy   301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCCGGAT  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2822 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCCGGAT 2881

Qy   361 AGCCTGTGGGATGGTGTGTTCAATTCCAGCATATAAATGGGGTCTCATCCTGCCCGGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2882 AGCCTGTGGGATGGTGTGTTCAATTCCAGCATATAAATGGGGTCTCGTCCTGCCCGGC 2941

Qy   421 ATCGTCATCCTCCTGTTACTGCATCATCATCCTCAAGCTGTCACACTCCAAGGGCCAC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2942 ATCGTCATCCTCCTGTTACTGCATCATCATCCTCAAGCTGTCACACTCCAAGGGCCAC 3001

Qy   481 CAGAGGCGCAAGGCCCTCAAGACGACAGTCATCCATCCTTCTTTGCCTGCTGG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3002 CAGAGGCGCAAGGCCCTCAAGACGACAGTCATCCATCCTTCTTTGCCTGCTCTG 3061

Qy   541 CTGCCCATATATGTGGGGATCAGCATCAGCATGCTCCTTTTGGGAGTCATCAAGCAA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3062 CTGCCCATATATGTGGGGATCAGCATCAGCATGCTCCTTTTGGGGGTCATCAAGCAA 3121

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGCCCTGCCTTC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3122 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGCCCTGCCTTC 3181

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
           ||||||||||||||||||||||||||||||
Db  3182 TTCCACTGTTGCCTGAACCCCATCCTCTAT 3211
```

RESULT 9
MMLESTRGN
LOCUS   MMLESTRGN   3770 bp   DNA   linear   ROD 02-APR-1997
DEFINITION M.musculus gene encoding leukocyte-derived seven transmembrane
            domain receptor, strain B6.
ACCESSION   X99581
VERSION   X99581.1 GI:1924959
KEYWORDS   lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Moepps,B., Frodl,R., Kessler,H. and Gierschik,P.
  TITLE   cDNA cloning and genomic organization of a leukocyte-derived seven
            transmembrane domain receptor (LESTR) from mouse: a murine
            homologue of the human HIV-1 entry cofactor fusin
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Moepps,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,
            Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK   Revised by [3]
REFERENCE   3 (bases 1 to 3770)
  AUTHORS   Moepps,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-APR-1997) B. Moepps, Universitaet Ulm,
            Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
COMMENT   On Apr 3, 1997 this sequence version replaced gi:1666646.
FEATURES   Location/Qualifiers
     source   1..3770
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="B6"
                     /db_xref="taxon:10090"
                     /haplotype="diploid"
                     /tissue_type="thymus"
                     /dev_stage="6-8 weeks old"
     exon   <1..110
                     /number=1
     gene   join(90..110,2306..3364)
                     /gene="lestr"
     CDS   join(90..110,2306..3364)
                     /gene="lestr"
                     /codon_start=1
                     /product="leukocyte-derived seven transmembrane domain
                     receptor"
                     /protein_id="CAA67893.1"
                     /db_xref="GI:1666647"
                     /db_xref="GOA:P70658"
                     /db_xref="SWISS-PROT:P70658"
                     /translation="MEPISVSIYTSDNYSEEVGSGGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GWIPALLLTIPDFIFADVSQGDISQGDRYICDRLYPDSLWMVFQFQHIMVGLVLP
                     GIVILSCYCIIISKLSHSKGHQKRXALKTVILILAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"
     misc_feature   105..2305
                     /gene="lestr"
                     /note="splice variant/intron"
     intron   111..2305
                     /gene="lestr"
     exon   2306..>3770
                     /number=2
     polyA_signal   3693..3698
ORIGIN

RESULT 10
AC145861
LOCUS   AC145861   175370 bp   DNA   linear   HTG 01-AUG-2003
DEFINITION Pan troglodytes chromosome UNK clone RP43-21D9, *** SEQUENCING IN
            PROGRESS ***, 39 unordered pieces.
ACCESSION   AC145861
VERSION   AC145861.1 GI:33386791
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Pan troglodytes (chimpanzee)
  ORGANISM Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 175370)
  AUTHORS   Wilson,R.K.
  TITLE   The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 175370)
  AUTHORS   Wilson,R.K.
  TITLE   Direct Submission

```
JOURNAL     Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT     -------------- Genome Center --------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information --------------
            Center project name: C_PT021D09

            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 39 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            *        1      5345: contig of 5345 bp in length
            *     5346      5445: gap of unknown length
            *     5446     17195: contig of 11750 bp in length
            *    17196     17295: gap of unknown length
            *    17296     33250: contig of 15955 bp in length
            *    33251     33350: gap of unknown length
            *    33351     45646: contig of 12296 bp in length
            *    45647     45746: gap of unknown length
            *    45747     61992: contig of 16246 bp in length
            *    61993     62092: gap of unknown length
            *    62093     63423: contig of 1331 bp in length
            *    63424     63523: gap of unknown length
            *    63524     64666: contig of 1143 bp in length
            *    64667     64766: gap of unknown length
            *    64767     65994: contig of 1228 bp in length
            *    65995     66094: gap of unknown length
            *    66095     67430: contig of 1336 bp in length
            *    67431     67530: gap of unknown length
            *    67531     68982: contig of 1452 bp in length
            *    68983     69082: gap of unknown length
            *    69083     70891: contig of 1809 bp in length
            *    70892     70991: gap of unknown length
            *    70992     72172: contig of 1181 bp in length
            *    72173     72272: gap of unknown length
            *    72273     73614: contig of 1342 bp in length
            *    73615     73714: gap of unknown length
            *    73715     75185: contig of 1471 bp in length
            *    75186     75285: gap of unknown length
            *    75286     76854: contig of 1569 bp in length
            *    76855     76954: gap of unknown length
            *    76955     78413: contig of 1459 bp in length
            *    78414     78513: gap of unknown length
            *    78514     79964: contig of 1451 bp in length
            *    79965     80064: gap of unknown length
            *    80065     81745: contig of 1681 bp in length
            *    81746     81845: gap of unknown length
            *    81846     83417: contig of 1572 bp in length
            *    83418     83517: gap of unknown length
            *    83518     84910: contig of 1393 bp in length
            *    84911     85010: gap of unknown length
            *    85011     86788: contig of 1778 bp in length
            *    86789     86888: gap of unknown length
            *    86889     88892: contig of 2004 bp in length
            *    88893     88992: gap of unknown length
            *    88993     90670: contig of 1678 bp in length
            *    90671     90770: gap of unknown length
            *    90771     93207: contig of 2437 bp in length
            *    93208     93307: gap of unknown length
            *    93308     96658: contig of 3351 bp in length
            *    96659     96758: gap of unknown length
            *    96759    100576: contig of 3818 bp in length
            *   100577    100676: gap of unknown length
            *   100677    104988: contig of 4312 bp in length
            *   104989    105088: gap of unknown length
            *   105089    108893: contig of 3805 bp in length
            *   108894    108993: gap of unknown length
            *   108994    112632: contig of 3639 bp in length
            *   112633    112732: gap of unknown length
            *   112733    116328: contig of 3596 bp in length
            *   116329    116428: gap of unknown length
            *   116429    120980: contig of 4552 bp in length
            *   120981    121080: gap of unknown length
            *   121081    126821: contig of 5741 bp in length
            *   126822    126921: gap of unknown length
            *   126922    131331: contig of 4410 bp in length
            *   131332    131431: gap of unknown length
            *   131432    138705: contig of 7274 bp in length
            *   138706    138805: gap of unknown length
            *   138806    145169: contig of 6364 bp in length
            *   145170    145269: gap of unknown length
            *   145270    150696: contig of 5427 bp in length
            *   150697    150796: gap of unknown length
            *   150797    159988: contig of 9192 bp in length
            *   159989    160088: gap of unknown length
            *   160089    167244: contig of 7156 bp in length
            *   167245    167344: gap of unknown length
            *   167345    175370: contig of 8026 bp in length.
FEATURES             Location/Qualifiers
     source          1..175370
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="UNK"
                     /clone="RP43-21D9"
     misc_feature    1..5345
                     /note="assembly_name:Contig100"
     misc_feature    5446..17195
                     /note="assembly_name:Contig101"
     misc_feature    17296..33250
                     /note="assembly_name:Contig102"
     misc_feature    33351..45646
                     /note="assembly_name:Contig103"
     misc_feature    45747..61992
                     /note="assembly_name:Contig104"
     misc_feature    62093..63423
                     /note="assembly_name:Contig16"
     misc_feature    63524..64666
                     /note="assembly_name:Contig3"
     misc_feature    64767..65994
                     /note="assembly_name:Contig45"
     misc_feature    66095..67430
                     /note="assembly_name:Contig48"
     misc_feature    67531..68982
                     /note="assembly_name:Contig5"
     misc_feature    69083..70891
                     /note="assembly_name:Contig56"
     misc_feature    70992..72172
                     /note="assembly_name:Contig60"
     misc_feature    72273..73614
                     /note="assembly_name:Contig67"
     misc_feature    73715..75185
                     /note="assembly_name:Contig68"
     misc_feature    75286..76854
                     /note="assembly_name:Contig69"
     misc_feature    76955..78413
                     /note="assembly_name:Contig73"
     misc_feature    78514..79964
                     /note="assembly_name:Contig75"
     misc_feature    80065..81745
                     /note="assembly_name:Contig77"
     misc_feature    81846..83417
                     /note="assembly_name:Contig79"
     misc_feature    83518..84910
                     /note="assembly_name:Contig80"
     misc_feature    85011..86788
                     /note="assembly_name:Contig81"
     misc_feature    86889..88892
```

```
misc_feature    /note="assembly_name:Contig82"
                88993. .90670
misc_feature    /note="assembly_name:Contig83"
                90771. .93207
misc_feature    /note="assembly_name:Contig84"
                93308. .96658
misc_feature    /note="assembly_name:Contig85"
                96759. .100576
misc_feature    /note="assembly_name:Contig86"
                100677. .104988
misc_feature    /note="assembly_name:Contig87"
                105089. .108893
misc_feature    /note="assembly_name:Contig88"
                108994. .112632
misc_feature    /note="assembly_name:Contig89"
                112733. .116328
misc_feature    /note="assembly_name:Contig90"
                116429. .120980
misc_feature    /note="assembly_name:Contig91"
                121081. .126821
misc_feature    /note="assembly_name:Contig92"
                126922. .131131
misc_feature    /note="assembly_name:Contig93"
                131432. .138705
misc_feature    /note="assembly_name:Contig94"
                138806. .145169
misc_feature    /note="assembly_name:Contig95"
                145270. .150696
misc_feature    /note="assembly_name:Contig96"
                150797. .159988
misc_feature    /note="assembly_name:Contig97"
                160089. .167244
misc_feature    /note="assembly_name:Contig98"
                167345. .175370
misc_feature    /note="assembly_name:Contig99"
```

```
Qy    481  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168678  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGG 168737

Qy    541  CTGCCATTATTATGTGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
           |||||||||||||||||||||||||||||||||||||||||| |||||  ||||||||||
Db 168738  CTGCCATTATTATGTGGGATCAGCATCGACTCCTTCATCCTTTCATCCTTTTGGGGTCATCAAGCAA 168797

Qy    601  GGATGTGACTTCGAGAGCATTGTGCACAAGTCGATCCATCACAGAGGCCCTCGGCCTTC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168798  GGATGTGACTTCGAGAGCATCGTGCCACAAGTCGATCCATCACACAGAGGCCCTCGGCCTTC 168857

Qy    661  TTCCACTGTTGCCTGACCCCATCCTCTAT 690
           ||||||||||||||||||||||||||||
Db 168858  TTCCACTGTTGCCTGAACCCCATCCTCTAT 168887
```

RESULT 11
AC147556

```
LOCUS       AC147556               185215 bp    DNA     linear   HTG 17-DEC-2003
DEFINITION  Mus musculus chromosome UNK clone RP23-480N14, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC147556
VERSION     AC147556.1  GI:39979584
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 185215)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 185215)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT
            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            ------------- Project Information -------------
            Center project name: M_BA0480N14
            ------------- Summary Statistics -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 194611 bases at least Q40
            Consensus quality: 196848 bases at least Q30
            Consensus quality: 198294 bases at least Q20
            Insert size: 182000; agarose-fp
            Insert size: 252319; sum-of-contigs
            Quality coverage: 13.47 in Q20 bases; agarose-fp
            Quality coverage: 9.68 in Q20 bases; sum-of-contigs
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 11 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *       1      1119: contig of 1119 bp in length
          *    1120      1219: gap of unknown length
          *    1220      2637: contig of 1418 bp in length
          *    2638      2737: gap of unknown length
```

ORIGIN

```
Query Match    98.8%;  Score 682;  DB 2;  Length 175370;
Best Local Similarity  99.3%;  Pred. No. 9e-161;
Matches  685;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy      1  CTGCACCTGTCAGTGGCTGACCTCCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db 168198  CTGCACCTGTCAGTGGCTGACCTCCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 168257

Qy     61  GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTCACACT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168258  GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTCACACT 168317

Qy    121  GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCCGGTACCTCGCC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168318  GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCCGGTACCTCGCT 168377

Qy    181  ATTGTCCACGGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168378  ATTGTCCACGGCCACCAACAGTCAGAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 168437

Qy    241  GTGGGCGTCTGGATCCCAGCGCCCTCCTGACTATACCTGACTTTCATCTTTGCCGACGTC 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168438  GTGGGCGTCTGGATCCCAGCGCCCTCCTGACTATACCTGACTTTCATCTTTGCCGACGTC 168497

Qy    301  AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACGGCCTTACCCCGGAT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168498  AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACGGCCTTACCCCGAT 168557

Qy    361  AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCATCCTGCCGGC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168558  AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCGTCCTGCCGGC 168617

Qy    421  ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168618  ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 168677
```

```
Db  91837 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGGCCTTTACCCCGAT 91896

Qy    361 AGCCTGTGGATGGTGGTGTTCAATTCAGCATATAATGGTGGTCTCATCCTGCCCGGC   420
Db  91897 AGCCTGTGGATGGTGGTGTTCAATTCAGCATATAATGGTGGTCTCGTCCTGCCCGGC 91956

Qy    421 ATGGTCATCCTCCTGTTACTGGCCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC   480
Db  91957 ATCGTCATCCTCCTCGTGTTACTGCCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 92016

Qy    481 CAGAAGCGCAAGGCCCTCAAGACGACAGTGCATCCTCATCCTCCTCAGCTTCTTGCCTGTGG   540
Db  92017 CAGAAGCGCAAGGCCCTCAAGACGACAGTGCATCCTCATCCTCCTCAGCTTTCTTGCCTGCTGG 92076

Qy    541 CTGCCATATTATGTGGGATCAGCATCGACTCCTCCTTTGGGAGTCATCAAGCAA   600
Db  92077 CTGCCATATTATGTGGGATCAGCATCGACTCCTCCTTTGGGGGTCATCAAGCAA 92135

Qy    601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGCCCTCGCCTTC   660
Db  92137 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGCCCCTCGCCTTC 92196

Qy    661 TTCCACTGTGCCTGAACCCCATCCTCTCAT   690
Db  92197 TTCCACTGTGCCTGAACCCCATCCTCTTAT 92226
```

RESULT 12
AF452185

```
FEATURES             Location/Qualifiers
     source          1..1050
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Holtzman"
                     /db_xref="taxon:10116"
                     /sex="female"
                     /tissue_type="whole brain"
     gene            1..1050
                     /gene="Cxcr4"
     CDS             1..1050
                     /gene="Cxcr4"
                     /codon_start=1
                     /product="chemokine receptor CXCR4"
                     /protein_id="AAL47855.1"
                     /db_xref="GI:17902281"
                     /translation="MEIYTSDNYSEEVGSGDYDSNKEPCFRDENENFNRIFLPTIYFI
                     IFLTGIVGNGLIVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFG
                     KFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIP
                     ALLLTIPDIIFADVSQGDGRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCI
                     IISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLEVIKQGCEFESV
                     VHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNMSRGSSLKILSKGKRGGH
                     SSVSTESESSSFHSS"
```

```
*  2738    5059: contig of 2322 bp in length
*  5060    5159: gap of unknown length
*  5160    6203: contig of 1044 bp in length
*  6204    6303: gap of unknown length
*  6304   11814: contig of 5511 bp in length
* 11815   11914: gap of unknown length
* 11915   18453: contig of 6539 bp in length
* 18454   18553: gap of unknown length
* 18554   33684: contig of 15131 bp in length
* 33685   33784: gap of unknown length
* 33785   57713: contig of 23929 bp in length
* 57714   57813: gap of unknown length
* 57814   84551: contig of 26738 bp in length
* 84552   84651: gap of unknown length
* 84652  122285: contig of 37634 bp in length
* 122286 122385: gap of unknown length
* 122386 185215: contig of 62830 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..185215
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP23-480N14"
     misc_feature    1..1119
                     /note="assembly_name:Contig51"
     misc_feature    1220..2637
                     /note="assembly_name:Contig103"
     misc_feature    2738..5059
                     /note="assembly_name:Contig122"
     misc_feature    5160..6203
                     /note="assembly_name:Contig124"
     misc_feature    6304..11814
                     /note="assembly_name:Contig126"
     misc_feature    11915..18453
                     /note="assembly_name:Contig127"
     misc_feature    18554..33684
                     /note="assembly_name:Contig128"
     misc_feature    33785..57713
                     /note="assembly_name:Contig129"
     misc_feature    57814..84551
                     /note="assembly_name:Contig130"
     misc_feature    84652..122285
                     /note="assembly_name:Contig131"
     misc_feature    122386..185215
                     /note="assembly_name:Contig132"
ORIGIN
```

```
Query Match           98.8%;  Score 682;  DB 2;  Length 185215;
Best Local Similarity  99.3%;  Pred. No. 9e-161;
Matches  685;  Conservative  0;  Mismatches  5;  Indels  0;  Gaps  0;

Qy     1 CTGCACCTGTCAGTCGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT   60
Db 91537 CTGCACCTGTCAGTCGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT 91596

Qy    61 GCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
Db 91597 GCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 91656

Qy   121 GTCAACCTCTCACAGCAGCGGTTCTCATCCTGGCCCTTCATCAGCGCTGGACCGGTACCTCGCC   180
Db 91657 GTCAACCTCTCACAGCAGCGGTTCTCATCCTGGCCCTTCATCAGCGCTGGACCGGTACCTCGCC 91716

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
Db 91717 ATTGTCCACGCCACCAACAGTCAGAGGCCCAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 91776

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC   300
Db 91777 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 91836

Qy   301 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGGCCTTTACCCCGAT   360
```

variation   429..430
                    /gene="Cxcr4"
                    /note="compared to Wistar strain sequence presented in
                    GenBank Accession Number U90610"
                    /replace="cg"

ORIGIN

Query Match   83.9%;   Score 578.6;   DB 10;   Length 1050;
Best Local Similarity 91.4%; Pred. No. 8e-135;
Matches 630; Conservative 0; Mismatches 44; Indels 15; Gaps 1;

Qy   1 CTGCACCTGCAGTGGCTGACCCCTCTTGTCATCACACTCCCCTTCTGGGCAGTGAT   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   223 CTGCACCTGTCCGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTGGAC 282

Qy   61 GCCATGGCTGACTCGTACTTTGGGAAATTTTTGTGTAAGGCGTCCATATCATCTACACT 120
            |||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   283 GCCATGGGCTGACTGGTACTTTGGGAAAATTTTTATGTAAGGCTGTGCATATCATCTACACC 342

Qy   121 GTCAACCTTCTACAGCAGCGGTTCATCCTGGCCTTCATCAGCCTGGACGGGTACCTCGCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   343 GTCAACCTTTACAGCAGTGTTCTTCCATCCTGGCCTTCATCAGCCTGGACCGGTACCTTGCC 402

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGAAAAGGCAGTCTAT 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   403 ATTGTCCACGCCACCAACAGCCAGAGGCCGAGGAAGTGCTGCTGCTGGCTGAAAAGGCCGTCTAT 462

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   463 GTGGGTGTCTGGATCCCCCGCCCCTCCTCCCTGACTATACCTGACATCATCTTCGCCGATGTC 522

Qy   301 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 360
            |||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   523 AGCCAGGGGAC------------GGCAGGTACATCTGTGACCGGCCTTTACCCCGAC 567

Qy   361 AGCCTGTGGATGGTGTTCAATTCCAGCATATAATGGTGGTCCATCCTGCCCGGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 AGCCTGTGGATGGTGGTGGTTCCCAGTTCCAGCACCATCATGGTGGTCCATCCTGCCGCGGC 627

Qy   421 ATCGTCATCCTCCCTGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCAC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628 ATCGTCATCCTCCTGTCCCTGTTACTGCATCATCATCCTCCAAGCTGTCACACTCCAAGGGCCAC 687

Qy   481 CAGAAGCGCGAAGGCCCTCAAGACGACACAGTCACTCATCCTCATCCTTTCTTTGCCTGCTGG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688 CAGAAGCGCGCAAGGCCCTCAAGACTACGGTCATCTCATCCTTATCCTGCTTTCTTTGCCTGCTGG 747

Qy   541 CTGCCATATTATGTGGGGATCAGCATGGCATCGACTCCTTGACTCCTTTTGGGAGTCATCAAGCAA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748 CTACCGTATTACGTGGGGGATCAGGCATGCATCGACTCCTTCATCCTTTTGGAGGTCATCAAGCAA 807

Qy   601 GGATGTGACTTGCCTGAGAGACCATTGTGCACAAGTGGATCCCATCACAGAGGCCCTCGCCTTC 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   808 GGATGTGAGTTGCAGTTCGAGAGCGTCGTGCACAAGTGGATCCCATCACGAGGCCCTCGCCTTC 867

Qy   661 TTCCACTGTGCCTGAACCCCATCCTCTA 689
            |||||||||||||||||||||||||||
Db   868 TTCCACTGTGCCTGAACCCCATCCTCTA 896

RESULT 13
AC122097
LOCUS   AC122097   190149 bp   DNA   linear   HTG 15-NOV-2002
DEFINITION Rattus norvegicus clone CH230-108A24, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC122097
VERSION   AC122097.4 GI:25008263
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

REFERENCE   1 (bases 1 to 190149)
AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Aisbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,J.,
            Davila.M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,S., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Pan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganca,R., Garcia,A., Garner,r., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow.H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen.H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,M., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,A., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Wilson,R., Wieczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,R., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith.H.O.,
            Weinstock,G. and Gibbs,R.A.
TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 190149)
AUTHORS   Worley.K.C.
TITLE   Direct Submission
JOURNAL   Submitted (22-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 190149)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 15, 2002 this sequence version replaced gi:23269619.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

-------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------- Project Information
    Center project name: GXCR
    Center clone name: CH230-108A24
-------- Summary Statistics
    Assembly program: Phrap; version 0.990329
    Consensus quality: 179135 bases at least Q40
    Consensus quality: 181484 bases at least Q30
    Consensus quality: 183011 bases at least Q20
    Estimated insert size: 182231; sum-of-contigs estimation
    Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
*   consists of 4 contigs. The true order of the pieces
*   is not known and their order in this sequence record is
*   arbitrary. Gaps between the contigs are represented as
*   runs of N, but the exact sizes of the gaps are unknown.
*   This record will be updated with the finished sequence
*   as soon as it is available and the accession number will
*   be preserved.
*   1   40067: contig of 40067 bp in length
*   40068   40167: gap of unknown length
*   40168   114660: contig of 74493 bp in length
*   114661   114760: gap of unknown length
*   114761   188186: contig of 73426 bp in length
*   188187   188286: gap of unknown length
*   188287   190149: contig of 1863 bp in length.

FEATURES   Location/Qualifiers
     source   1..190149
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-108A24"

ORIGIN

Query Match   83.9%; Score 578.6; DB 2; Length 190149;
Best Local Similarity 91.4%; Pred. 9.7e-135;
Matches 630; Conservative   0; Mismatches 44; Indels 15; Gaps   1;

Qy   1 CTGCACCTGCTCAGTGGCTGACCCTCCTTCGTCATCACACTCCCCTTCTCGGGCAGTGAT 60
          ||||||||||||| || |||||||||||||| |||||||||||||| ||||||||||||||
Db 20878 CTGCACCTGTCCGTGGCTGACCTCCTTCTTTGTCATCACACTCCCCTTCTGGGCAGTGGAC 20937

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCGTGTCCATATCATCTACACT 120
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 20938 GCCATGGCTGACTGGTACTTTTGGGAAATTTTTATGTAAGGCGTGTCATATCATCTCACACC 20997

Qy   121 GTCAACCTTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 20998 GTCAACCTTTACAGCAGTGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTTGCC 21057

Qy   181 ATTGTCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
          ||||||||||||| || |||||||||||||||||||||||||||||||||||||||||
Db 21058 ATTGTCCACGCCACCAACAGCCAGAGCCGAGGAAGCTGCTGGCTGAAAAAGGCCGTCTAT 21117

Qy   241 GTGGGCGTCTGGATCCCAGCCCTTCCCGCCCTCCTTCCTGACTTACCTGACTTCATCTTT 300
          ||||||||||||||||||||||||||||| || ||||| ||||||||||||||||||||
Db 21118 GTGGGTGTCTGGATCCCGCCGCCCCTTCCCGCCCTCCTTCCTGACATCATCTTCTCCGGATGTC 21177

Qy   301 AGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGGACGCCCTTTACCCCGAT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 21178 AGCCAGGGGAC-----------GGCAGGTACATCTGTGGACCGCCCTTTACCCCGAC 21222

Qy   361 AGCCTGGGATGGTGTGTTTCAATTCCAGCATATAATGTGGGTCTCATCCTGCCCGGC 420
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 21223 AGCCTGGGGATGGTGTGTTCCAGTTCCAGCACATCATGGGGTCTCATCCTGCCCGGGC 21282

Qy   421 ATCGTCATCCTCCGTGTTACTGCATCATCATCTTCTCAAGTGTCACACTCCAAGGGCCAC 480
          |||||||||||| || |||||||||||||||||||||||||||||||||||||||||||
Db 21283 ATCGTCATCCTGTCCGTTACTGCATCATCATCATCCCAAGCTGTCACACTCCAAGGGCCAC 21342

Qy   481 CAGAAGCGCAAGGCCCTCAAGACGACAGTGACTCATCCTCATCCTTTCTTTGCCTGCTGG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 21343 CAGAAGCGCAAGGCCCTCAAGACTCAGTGACTCATCCTCATCCTTTCTTTTGCCTGCTGG 21402

Qy   541 CTGCCATATTATGTGGGATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 21403 CTACCGTATTACGTGGGGATCAGCATCGATCATCCTTCATCCTTTGGAGGTCATCAAGCAA 21462

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTGACAAGTGGATCCATCAGAGGCCCTCGCCTTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 21463 GGATGTGAGTTCGAGAGCGTCGTGCACAAGTGGGATCCATCACCGGAGGCCCTCGCCTTC 21522

Qy   661 TTCCACGTGTGCCTGAACCCATCCTCTA 689
          |||||||||||||||||||||||||||||
Db 21523 TTCCACGTGTGCCTGAACCCCATCCTCTA 21551

RESULT 14
AC107097/c
LOCUS   AC107097   240931 bp   DNA   linear   HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-175010, *** SEQUENCING IN PROGRESS
            ***
ACCESSION   AC107097
VERSION   AC107097.5 GI:30580624
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 240931)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen.H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Burch,J., Burrell,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,J.,
            Davila.M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,P., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,R., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,R., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,H., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.L.,

Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,C., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Wilson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 240931)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (16-JAN-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 240931)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
COMMENT   On May 13, 2003 this sequence version replaced gi:23115278.
           The sequence in this assembly is a combination of BAC based reads
           and whole genome shotgun sequencing reads assembled using Atlas
           (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
           in the feature table below represents a scaffold in the Atlas
           assembly (a 'contig-scaffold'). Within each contig-scaffold,
           individual sequence contigs are ordered and oriented, and separated
           by sized gaps filled with Ns to the estimated size. The sequence
           may extend beyond the ends of the clone and there may be sequence
           contigs within a contig-scaffold that consist entirely of whole
           genome shotgun sequence reads. Both end sequences and whole genome
           shotgun sequence only contigs will be indicated in the feature
           table.
           --------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
           --------- Project Information
              Center project name: GKEK
              Center clone name: CH230-175O10
           --------- Summary Statistics
              Assembly program: Atlas 3.0;
              Consensus quality: 216793 bases at least Q40
              Consensus quality: 219254 bases at least Q30
              Consensus quality: 221224 bases at least Q20
              Estimated insert size: 228277; sum-of-contigs estimation
              Quality coverage: 6x; sum-of-contigs estimation
           * NOTE: Estimated insert size may differ from sequence length
           *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
           * NOTE: This is a 'working draft' sequence. It currently
           *   consists of 1 contigs. Gaps between the contigs
           *   are represented as runs of N. The order of the pieces
           *   is believed to be correct as given, however the sizes
           *   of the gaps between them are based on estimates that have
           *   provided by the submittor.
           *   This sequence will be replaced
           *   by the finished sequence as soon as it is available and
           *   the accession number will be preserved.
           *   1 240931: contig of 240931 bp in length.
FEATURES   Location/Qualifiers
     source   1..240931
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-175O10"
     misc_feature   1..1249
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature   complement(5161..5587)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH357332"
     misc_feature   complement(232579..233220)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH357330"
     misc_feature   235468..237558
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature   239450..240931
                     /note="wgs_end_extension
                     clone_end:T7"

ORIGIN

Query Match   83.9%;   Score 578.6;   DB 2;   Length 240931;
Best Local Similarity 91.4%;   Pred. No. 9.8e-135;
Matches 630;   Conservative   0;   Mismatches 44;   Indels 15;   Gaps 1;

```
Qy       1 CTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACAACACTCCCCTTCTGGGCAGTGAT     60
           ||||||||||| |||||||||||||||  |||||||||  |||||||||||||||||||
Db  237401 CTGCACCTGTCCGTGGCTGACCTCCTCTTGTCATCACACCACTCCCCTTCTGGGCAGTGGAC 237342

Qy      61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTCACACT    120
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  237341 GCCATGGCTGACTGGTACTTTGGGAAATTTTTATGTAAGGCTGTCCATATCATCTCACACC 237282

Qy     121 GTCAACCTCTACAGACGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCC    180
           |||||||  |||| ||  |||||||||||||||||||||||||||||||| ||  ||||
Db  237281 GTCAACCTTTCAGCAGTGTTCTCATCCTGGCCTTCATCAGCCTGACCGGCCTACCTTGCC 237222

Qy     181 ATTGTCACGCCACCACCAGTCAAAGCCAAGGAAACTGCTGGCTGGCTGCTGAAAAGCCAGTCTAT    240
           |||||| ||||||||||||  |  |||| |||| |||||||||||||||||||||||  |||||
Db  237221 ATTGTCCACGCCACCACCAGCCAGGAGGCCGAGGAAGCTGCTGGCTGGCTGCTGAAAAGCCGGTCTAT 237162

Qy     241 GTGGGGGTCTGGATGCCCAGCCCTCCTGACTATACCTGACTTCATCTTGCCGACGTC    300
           ||||| ||||||||||| ||||||||||||||  |||||| | ||||||||||| |||
Db  237161 GTGGGGTGTCTGGATGCCCGGCCCTCCTGACTATCCTGACTATCATCTTGCCGACGATGTC 237102

Qy     301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGGCCTTTACCCCGAT    360
           |||||||||||            ------GGCAGGTACATCTGTGACCGGCCTTTACCCCGAC
Db  237101 AGCCAGGGGGAC-------------GGCAGGTACATCTGTGACCGGCCTTTACCCCGAC 237057

Qy     361 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATATGGTGGTCTCATCCTGCCCGGC    420
           ||||||||||||||||||||| || ||||||||| | |||||| |||||||||||||||
Db  237056 AGCCTGTGGATGGTGGTGTTCCAGTTCCAGCACATCATGGTGGGTCTCATCCTGCCCGGGC 236997

Qy     421 ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC    480
           ||||||||||||||||  |  |||||||||||||| ||||||||||||||||||||||||
Db  236996 ATCGTCATCCTCCTGTCCTCGTTACTGCATCATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 236937

Qy     481 CAGAAGCGCAAGGCCCTCAAGACGAGCGAGCTGTGCACAGTGGATCCTCATCCTCATCCTCAGCTTTCTTTGCCTGTGG    540
           |||||||||||||||||||||||  ||  |  ||  ||||||  |||||||||  ||||||||||
Db  236936 CAGAAGCGCAAGGCCCTCAAGACTACGGTCATCATCCTTATCCTGGCTTTCTTTGCTCGTGG 236877

Qy     541 CTGCCATATTATGTGGGGATCAGCATCGACTTCCTTCATCCTTTGGGAGTCATCAAGCAA    600
           ||  ||||||| |||||||||||||||  |||||||   |||||||| ||||||||||
Db  236876 CTACCGTATTACGTGGGGATCAGCATCGAATTCCTTTTGGAGGTCATCAAGCAA 236817

Qy     601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGAGCCCTCGCCTTC    660
           ||||||||| |||||||| | ||||||||||||||| |||||||||||||||||||||||
Db  236816 GGATGTGAGTTCGAGAGCGTCGTGCACAAGTGGATCGATGCACAAGTGGATCCCATCACAGAGGAGCCCTCGCCTTC 236757

Qy     661 TTCCACTGTTGCTGAACCCCATCCTCTA    689
```

```
Db    236756  TTCCACTGTTGCCTGAACCCATCCTCTA 236728
                |||||||||||||||||||||||||||

RESULT 15
RNU90610
LOCUS       RNU90610        1050 bp    mRNA    linear    ROD 26-MAR-1997
DEFINITION  Rattus norvegicus CXC chemokine receptor (CXCR4) mRNA, complete
            cds.
ACCESSION   U90610
VERSION     U90610.1  GI:1906612
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1050)
  AUTHORS   Harrison,J.K. and Salafranca,M.N.
  TITLE     Molecular cloning of rat CXCR4
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1050)
  AUTHORS   Harrison,J.K. and Salafranca,M.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-FEB-1997) Pharmacology and Therapeutics, University
            of Florida, 1600 SW Archer Rd., Gainesville, FL 32610-0267, USA
FEATURES             Location/Qualifiers
     source          1..1050
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /tissue_type="spleen"
     gene            1..1050
                     /gene="CXCR4"
     CDS             1..1050
                     /gene="CXCR4"
                     /note="G-protein coupled receptor"
                     /codon_start=1
                     /product="CXC chemokine receptor"
                     /protein_id="AAB50408.1"
                     /db_xref="GI:1906613"
                     /translation="MEIYTSDNYSEEVGSGDYDSNKEPCFRDENENFNRIFLPTIYFI
                     IFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFG
                     KFLCKAVHIIYTVNLYSSVLILAPISLDRYLAIVHATNSQSARKLLAEKAVVGVWIP
                     ALLLTIPDIIFADVSQGDGRYICDRLYFDSLMVVFQFQHIMVGLILPGIVILSCYCI
                     IISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLEVIKQGCEFESV
                     VHKWISITEBALAPFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGH
                     SSVSTESESSFHSS"
ORIGIN

Query Match      83.4%;   Score 575.4;  DB 10;  Length 1050;
Best Local Similarity   91.1%;   Pred. No. 5.1e-134;
Matches  628;   Conservative   0;   Mismatches  46;   Indels  15;   Gaps   1;

Qy    1   CTGCACCTGTCAGTGGCTGACCTCCTTCTTTGTCATCACACTCCCTTCTCGGGCAGTTGAT  60
          ||||||||||||||||||||||||||||||||||||||||| ||||||| |||||||||||
Db   223  CTGCACCTGTCCGTGGCTGACCTCCTTCTTTGTCATCACACTCCCTTCTCGGGCAGTTGGAC 282

Qy    61  GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTTACACT 120
          |||||||||||||||||||||||||||||||  |||||||||||| |||||||| |||||
Db   283  GCCATGGCTGACTGGTACTTTGGGAAATTTTTATGTAAGGCTGTGCATATCATCTTACACC 342

Qy   121  GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCTCATCAGCCTGGACCGGTACCTCGCC 180
          |||||| ||||||||||||||||||||||||||||| ||||||||||||||||||| |||||
Db   343  GTCAACCTTTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGCTACCTTGCC 402

Qy   181  ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGGCTGAAAAGGCAGTCTAT 240
          ||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||
Db   403  ATTGTCCACGCCACCAACAGCCAGAGCGCGAGGAAGCTGCTGGCTGAAAAGGCCGTCTAT 462

Qy   241  GTGGGCGTCGTCGGATCCCAGCCCTCCTTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
          |||||| ||||||||||||||||||||| |||||| ||||||||||||||| ||||||| |||
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    May 17, 2004, 03:46:58 ; Search time 295.894 Seconds
                              (without alignments)
                              9906.453 Million cell updates/sec

Title:          US-09-367-052-3
Perfect score:  690
Sequence:       1 ctgcacctgtcagtggctga.........gcctgaacccatcctctat 690

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:     6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      N_Geneseq_29Jan04:*
                1:  geneseqn1980s:*
                2:  geneseqn1990s:*
                3:  geneseqn2000s:*
                4:  geneseqn2001as:*
                5:  geneseqn2001bs:*
                6:  geneseqn2002s:*
                7:  geneseqn2003as:*
                8:  geneseqn2003bs:*
                9:  geneseqn2003cs:*
                10: geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	690	100.0	1758	2	AAZ27611	Aaz27611 Mouse CXC
2	690	100.0	1877	2	AAV46370	Aav46370 Nucleic a
3	575.4	83.4	1050	9	ADB58421	Adb58421 Toxicity-
4	575.4	83.4	1050	9	ADB53005	Adb53005 Primary r
5	517.2	75.0	1059	3	AAZ38553	Aaz38553 Human CXC
6	517.2	75.0	1059	6	ABQ74938	Abq74938 Human CXC
7	517.2	75.0	1059	6	AAH99952	Aah99952 CXCR4 enc
8	517.2	75.0	1059	6	ACC72707	Acc72707 Human can
9	517.2	75.0	1225	6	ABZ35630	Abz35630 Human gen
10	517.2	75.0	1225	7	ACA56637	Aca56637 Human sig
11	517.2	75.0	1317	2	AAV18357	Aav18357 Human RM3
12	517.2	75.0	1317	3	AAA91726	Aaa91726 Human 7TM
13	517.2	75.0	1317	6	ABK54255	Abk54255 Human 7 t
14	517.2	75.0	1582	6	ABS64793	Abs64793 CXCR4 rec
15	517.2	75.0	1588	2	AAZ27610	Aaz27610 Human CXC
16	517.2	75.0	1664	3	AAB61656	Aab61656 Human CXC
17	517.2	75.0	1664	3	AAZ40014	Aaz40014 CXCR4 cod
18	517.2	75.0	1670	6	ABK83803	Abk83803 Human cDN
19	517.2	75.0	1670	6	ABN95645	Abn95645 Gene #214
20	517.2	75.0	1670	7	ACC46765	Acc46765 Human COP
21	517.2	75.0	1679	4	AAD12801	Aad12801 Human neu
22	517.2	75.0	1679	4	AAI65467	Aai65467 Nucleotid
23	517.2	75.0	1679	6	ABS53992	Abs53992 DNA encod
24	517.2	75.0	1679	6	ABS53750	Abs53750 DNA encod
25	517.2	75.0	1679	6	ABV78039	Abv78039 Hypoxia-r
26	517.2	75.0	1679	7	ABX08779	Abx08779 Angiogene
27	517.2	75.0	1679	7	ABX74454	Abx74454 Human cDN
28	517.2	75.0	1679	7	ABZ68886	Abz68886 Nucleotid
29	517.2	75.0	1679	7	ABZ42642	Abz42642 Human CXC
30	517.2	75.0	1679	9	ADC98645	Adc98645 Human CXC
31	517.2	75.0	1711	9	ABZ35348	Abz35348 Human gen
32	517.2	75.0	1711	9	ADB47320	Adb47320 Human cDN
33	517.2	75.0	1737	2	AAQ80521	Aaq80521 Human mon
34	517.2	75.0	1737	2	AAQ99007	Aaq99007 Chemokine
35	517.2	75.0	1944	2	AAX15882	Aax15882 cDNA enco
36	517.2	75.0	5161	3	AAA34774	Aaa34774 Human ade
37	517.2	75.0	5161	3	AAF20896	Aaf20896 Human CCR
38	517.2	75.0	5161	6	AAH99951	Aah99951 CXCR4 gen
39	517.2	75.0	5161	7	ABZ96590	Abz96590 Human CCR
40	517.2	75.0	5161	7	ACA64841	Aca64841 Human CXC
41	515.6	74.7	1102	7	ACA64750	Aca64750 Chemokine
42	515.6	74.7	1317	9	AAQ66179	Aaq66179 Seven tra
43	515.6	74.7	8747	9	ADD67548	Add67548 Human Ly1
44	514	74.5	1737	2	AAQ29506	Aaq29506 New plate
45	512.4	74.3	1059	5	ABI97982	Abi97982 Non-endog

ALIGNMENTS

```
RESULT 1
AAZ27611
ID   AAZ27611 standard; DNA; 1758 BP.
XX
AC   AAZ27611;
XX
DT   16-DEC-1999  (first entry)
XX
DE   Mouse CXCR4 coding sequence.
XX
KW   CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;
KW   tissue repairing agent; vascularisation; ss.
XX
OS   Mus sp.
XX
PN   WO994528-A1.
XX
PD   30-SEP-1999.
XX
PF   23-MAR-1999;   99WO-JP001448.
XX
PR   24-MAR-1998;   98JP-0095448.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
PA   (KISH/) KISHIMOTO T.
XX
PI   Kishimoto T,  Nagasawa T,  Tachibana K;
XX
DR   WPI; 1999-591042/50.
DR   P-PSDB; AAY39994.
XX
PT   CXCR4-potentiating agents and methods useful for inhibiting
PT   neovascularization, and treating solid cancers.
XX
PS   Disclosure; Page 50-51; 63pp; Japanese.
XX
CC   This sequence encodes the mouse CXCR4 protein. The invention relates to
CC   remedies inhibiting neovascularisation, remedies for solid cancer,
CC   remedies for diseases pathologically caused by neovascularisation and
CC   tissue repairing agents containing as the active ingredient a substance
CC   capable of potentiating CXCR4. Based on a finding that vascularisation is
CC   inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC   remedies inhibiting vascularisation which contain as the active
CC   ingredient a substance capable of potentiating CXCR4, remedies for solid
CC   cancer, remedies for diseases pathologically caused by neovascularisation
CC   and tissue repairing agents containing as the active ingredient a
```

Query Match   100.0%; Score 690; DB 2; Length 1758;
Best Local Similarity   100.0%; Pred. No. 6.8e-186;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  238 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT  297

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  298 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACT  357

Qy  121 GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  358 GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  417

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGTGGCTGAAAAGGCAGTCTAT  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  418 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGTGGCTGAAAAGGCAGTCTAT  477

Qy  241 GTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  478 GTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  537

Qy  301 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCCGTGACGCCCTTTACCCCGAT  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  538 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCCGTGACGCCCTTTACCCCGAT  597

Qy  361 CTGCCATTATGTGGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAA  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  598 CTGCCATTATGTGGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAA  657

Qy  421 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCCAGAGGCCCTGCCTTC  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  658 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCCAGAGGCCCTGCCTTC  717

Qy  661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
        ||||||||||||||||||||||||||||||
Db  898 TTCCACTGTTGCCTGAACCCCATCCTCTAT  927
```

RESULT 2
AAV46370
ID   AAV46370 standard; cDNA to mRNA; 1877 BP.
XX
AC   AAV46370;
XX
DT   20-NOV-1998 (first entry)
XX
DE   Nucleic acid encoding a murine CXC chemokine receptor.
XX
KW   Mouse; CXC chemokine receptor; pre-B cell line DW34;
KW   CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1; HIV infection;
KW   screening; inhibitor; AIDS; ds.
XX
OS   Mus sp.

XX
FH   Key   Location/Qualifiers
FT   CDS   120..1199
FT   /*tag= a
XX
PN   WO9835035-A1.
XX
PD   13-AUG-1998.
XX
PF   07-FEB-1997; 97WO-JP000299.
XX
PR   07-FEB-1997; 97WO-JP000299.
XX
PA   (SHIO) SHIONOGI & CO LTD.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;
PI   Nakajima T, Yoshie O;
XX
DR   WPI; 1998-447232/38.
DR   P-PSDB; AAW64778.
XX
PT   Mouse CXC chemokine receptor binding to PBSF/SDF-1 pre-B cell stimulatory
PT   factor - is useful for screening of potential HIV infection and AIDS
PT   inhibitors.
XX
PS   Claim 3; Page 39-42; 76pp; Japanese.
XX
CC   The present sequence encodes a murine CXC chemokine receptor which binds
CC   to the mouse CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1. The
CC   nucleic acid is isolated from mouse pre-B cell line DW34. The receptor
CC   and cells expressing it can be used in the study and mapping of the
CC   mechanism of HIV infection and in screening of potential inhibitors of
CC   HIV infection and the development of AIDS
XX
SQ   Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T; 0 U; 0 Other;

Query Match   100.0%; Score 690; DB 2; Length 1877;
Best Local Similarity   100.0%; Pred. No. 7e-186;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  357 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT  416

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  417 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACT  476

Qy  121 GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  477 GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  536

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGTGGCTGAAAAGGCAGTCTAT  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  537 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGTGGCTGAAAAGGCAGTCTAT  596

Qy  241 GTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  597 GTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  656

Qy  301 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCCGTGACGCCCTTTACCCCGAT  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  657 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCCGTGACGCCCTTTACCCCGAT  716

Qy  361 ATCGTGTGGATGGTGGTTTCAATTCCAGCATATATGGGTGGTCTCATCCTGCCCGGC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  717 ATCGTGTGGATGGTGGTTTCAATTCCAGCATATATGGGTGGTCTCATCCTGCCCGGC  776

Qy  421 ATCGTCATCCTCCTCGTTACTGCATCATCATCCTCTAAGCTGTCACACTCCAAGGGCCAC  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  777 ATCGTCATCCTCCTCGTTACTGCATCATCATCCTCTAAGCTGTCACACTCCAAGGGCCAC  836

Qy  481 CAGAAGCGCAAGGCCCTCAAGACGCCCTCAAGAGACGA  540
        ||||||||||||||||||||||||||||||||||||||
Db  481 CAGAAGCGCAAGGCCCTCAAGACGCCCTCAAGAGACGA  540
```

```
SQ   Sequence 1050 BP; 231 A; 313 C; 247 G; 259 T; 0 U; 0 Other;

     Query Match        83.4%;   Score 575.4;  DB 9;  Length 1050;
     Best Local Similarity  91.1%;   Pred. No. 2.3e-153;
     Matches 628;  Conservative   0;  Mismatches 46;  Indels 15;  Gaps   1;

Qy      1 CTGGCACCTGTCAGTGCGTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
          ||||||||||| |||||| |||||||||||||||||||||||||||||||||||||| |||
Db    223 CTGCACCTGTCCGGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTGGAC 282

Qy     61 GCCATGGCTGACTGACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
          |||||||||||||||||||||||||||| ||||||||||||||||| ||| ||||||
Db    283 GCCATGGCTGACTGACTTTGGGAAATTTTATGTAAGGCTGTCATATCATCTACACC 342

Qy    121 GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    343 GTCAACCTTTACAGCAGCAGTGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTTGCC 402

Qy    181 ATTGTCACGGCCACCAACAGTCAAAGGCCAAGGAAACTGTGGCTGAAAAGGCAGTCTAT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    403 ATTGTCCAGGCCACCAACAGCCAGGCGCGGAGGAAGCTGCTGGCTGAAAAGGCCGTCTAT 462

Qy    241 GTGGGCGTTGGATCCCAGCCCTCCTGACTTCATCATCTTTGCCGACGTC 300
          |||||||||||||||||||||||||||||| ||||||||||||||||||
Db    463 GTGGGTGTCGGATCCCGCCCTCCTGACTATCCCGACATCATCTTCGCCGATGTC 522

Qy    301 AGCCAGGGGACATCAGTCAGGGGGATGACGCAGGTACATCTGTGACCGCCGCCTTACCCCGAT 360
          |||||||||||                     ||||||||||||||||||||||||||||||
Db    523 AGCCAGGGGGAC---------GGCAGGTACATCTGTGACCGCGGCTTTACCCCGAC 567

Qy    361 AGCCTGTGGATGGTGCGTGTTCAATTCCAGCATATAATGGTGGGTCGTCACACTCCAAGGGCCAC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    568 AGCCTGTGGATGGTGGTGTTCCAGTTGCATCATCATCTCCAAGCGTGTCACACTCCAAGGGCCAC 627

Qy    421 ATCGTCATCCTCCTCGTACTGTCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628 ATCGTCATCCTGTCCTGTTACTGCATCATCATCTCCAAGCGTGTCACACTCCAAGGGCCAC 687

Qy    481 CAGAAGCGCAAGGCCCTCAAGACGACCAGTCATTCTTGCCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||
Db    688 CAGAAGCGCAAGGCCCTCAAGACTACGGTCATCCTTGCCTCTGG 747

Qy    541 CTGCCATATATGTGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748 CTACCGTATTACGTGGGGATCAGCATCGATTCCTTCATCCTTTTTGGAGGTCATCAAGCAA 807

Qy    601 GGATGTGACTTGAGAGCATTGTGCCACAAGTGGATCTCCATCACAGAGAGCCCTGCCTTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808 GGATGTGAGTTGAGAGCGTGTGCCACAAGTGGATCTCCATCACGGAGAGCCCTGCCTTC 867

Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTA 689
          |||||||||||||||||||||||||||||
Db    868 TTCCACTGTTGCCTGAACCCCATCCTCTA 896
```

```
RESULT 4
ADB53005
ID   ADB53005 standard; DNA; 1050 BP.
XX
AC   ADB53005;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3547.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
```

Left column:

```
Db    837 CAGAAGCGCAAGGCCCTCAAGACGACGAGCATCGACAGCGACAGTCATCCTCAGCCTCTTCTTTGCCTGCTGG 896
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 CTGCCATATATGTGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    897 CTGCCCATATATGTGGGGATCAGCATCGATTCCTTCATCCTTTTTGGGAGTCATCAAGCAA 956

Qy    601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGCCCTGCCTTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    957 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGCCCTCGCCTTC 1016

Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
          ||||||||||||||||||||||||||||||
Db   1017 TTCCACTGTTGCCTGAACCCCATCCTCTAT 1046
```

```
RESULT 3
ADB58421
ID   ADB58421 standard; DNA; 1050 BP.
XX
AC   ADB58421;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Toxicity-related gene, SEQ ID 3447.
XX
KW   Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW   drug screening; toxicity assay; ds.
XX
OS   Unidentified.
XX
PN   WO2003064624-A2.
XX
PD   07-AUG-2003.
XX
PF   31-JAN-2003; 2003WO-US003194.
XX
PR   31-JAN-2003; 2002US-00060087.
PR   15-MAR-2002; 2002US-0364045P.
PR   15-MAR-2002; 2002US-0364055P.
PR   30-DEC-2002; 2002US-0436643P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
DR   WPI; 2003-689530/65.
XX
PT   Predicting a toxic effect of a compound, useful in identifying toxicity
PT   markers in liver tissues or cells for drug screening and toxicity assays,
PT   comprises preparing gene expression profile of tissue or cells exposed to
PT   the compound.
XX
PS   Claim 1; SEQ ID NO 3447; 1156pp; English.
XX
CC   The present invention relates to a method for predicting a toxic effect
CC   of a compound. The method comprises preparing a gene expression profile
CC   of a tissue or cell sample exposed to the compound, and comparing the
CC   gene expression profile to a database comprising SEQ ID 1-4925, where
CC   differential expression of the gene indicates at least one toxic effect.
CC   The method is useful for predicting at least one toxic effect of a
CC   compound, predicting hepatotoxicity or the progression of a toxic effect
CC   of a compound, identifying an agent that modulates the onset or
CC   progression of a toxic response, predicting the cellular pathways that a
CC   compound modulates in a cell, and identifying an agent that modulates at
CC   least one activity of a protein. The method and compositions of the
CC   present invention using a database of genes having liver toxin-induced
CC   differential expression, are useful in identifying toxicity markers in
CC   liver tissues or cells for drug screening and toxicity assays. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
```

PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
PR   19-APR-2002; 2002US-0373601P.
PR   19-APR-2002; 2002US-0373602P.
PR   22-APR-2002; 2002US-0374139P.
PR   08-MAY-2002; 2002US-0378370P.
PR   09-MAY-2002; 2002US-0378652P.
PR   09-MAY-2002; 2002US-0378653P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-JUL-2002; 2002US-0394230P.
PR   09-JUL-2002; 2002US-0394253P.
PR   04-SEP-2002; 2002US-0407688P.
PR   28-JAN-2003; 2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3547; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 1050 BP; 231 A; 313 C; 247 G; 259 T; 0 U; 0 Other;

Query Match   83.4%; Score 575.4; DB 9; Length 1050;
Best Local Similarity   91.1%; Pred. No. 2.3e-153;
Matches   628; Conservative   0; Mismatches 46; Indels 15; Gaps   1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   223 CTGCACCTGTCCGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTGAC 282

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGCTGTCCATATCATCTACACT 120
          |||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Db   283 GCCATGGCTGACTGGTACTTTGGGAAATTTTATGTAAGGCTGCTGCATATCATCTACACC 342

Qy   121 GTCAACCTGCTACAGCAGCCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   343 GTCAACCTTTACAGCAGTGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGCTACCTTGCC 402

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   403 ATTGTCCACGCCACCAACAGCCCAGAGCGCGAGGAAGCTGCTGGCTGCTGAAAAGGCCGTCTAT 462

                                                    Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
                                                              ||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
                                                    Db   463 GTGGGTGTCTGGATCCCGCCGCCTCCTCCTGACTATCCTCCTGACATCATCTTCGCCGACGTC 522

                                                    Qy   301 AGCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGGCCTTTACCCGGAT 360
                                                              ||||||||| |||||||||||||||||
                                                    Db   523 AGCAGGTACATCGTGACCGGCCCTTTACCCGGAC 567

                                                    Qy   361 AGCCTGTGTGGATGTGGTGTTCAATTCCAGCATATAATGGTGGGTCCATCCTGCCCGGC 420
                                                              |||||||||||||||||||||||||||||||||||||||| |||||||||||
                                                    Db   568 AGCCTGTGTGGATGTGGTGTTCCAGTTCCAGCACATCATGGTGGGCTCATCCTGCCCGGGC 627

                                                    Qy   421 ATCGTCATCCTCCTGTTACTGCATCATCATCTCTTAAGCTGTCACACTCCAAGGGCCAC 480
                                                              ||||||||||| |||| |||||| |||| ||||||||||||||||||||||||
                                                    Db   628 ATCGTCATCCTCCTGTCCCTGTTACTGCATCATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 687

                                                    Qy   481 CAGAAGCGCAAGGCCCTCAAGACGACAGTGCATCCTCCAGCTTTCTTTGCCTGCTGG 540
                                                              |||||||||||||||||||||||||||||||||||||||||| |||||| |||||||
                                                    Db   688 CAGAAGCGCAAGGCCCTCAAGACTACGGTCATCCTTCATCCTTATCCTGGCTTTCTTTGCCTGCTGG 747

                                                    Qy   541 CTGCCATATTATGTGGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAA 600
                                                              |||||||| ||||||| |||||||||||||| |||||||||||||||||||||
                                                    Db   748 CTACCGTATTACGTGGGGATCAGCATCGATTCCTTCATCCTTTGGAGGTCATCAAGCAA 807

                                                    Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTTCCATCACAGAGGCCCTCGCCTTC 660
                                                              ||||||||||||||||||||||||| ||||||||||||| |||||||||||||||
                                                    Db   808 GGATGTGAGTTCGAGAGCGTCTGCGCACAAGTGGATCTCCATCACGCGGAGGCCCTCGCCTTC 867

                                                    Qy   661 TTCCACTGTTGCCTGACCCCATCCTCTA 689
                                                              |||||||||||||||||| |||||
                                                    Db   868 TTCCACTGTTGCCTGACCCCATCCTCTA 896

RESULT 5
AAZ38553
ID   AAZ38553 standard; cDNA; 1059 BP.
XX
AC   AAZ38553;
XX
DT   22-FEB-2000 (first entry)
XX
DE   Human CXC-chemokine receptor-4 (CXCR-4) cDNA.
XX
KW   CXC-chemokine receptor-4; CXCR-4; stromal cell derived factor-1; SDF-1;
KW   ligand; chemotaxis; inflammation; G-protein-coupled receptor;
KW   signal transduction; CD4-independent; HIV-1; infection; proliferation;
KW   transformation; tumorigenesis; cancer; melanoma; overexpression; brain;
KW   breast; colon; lung; melanoma; glioblastoma; inhibition; growth arrest;
KW   diagnosis; prognosis; marker; proliferative disorders; antisense;
KW   therapy; treatment; premalignant condition; hypertrophy;
KW   degenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   growth deficiency; hypoproliferative disorder; physical trauma; lesion;
KW   ischaemia; wound; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1059
FT   /*tag= a
FT   /product= "Human CXCR-4"
XX
PN   WO9950461-A1.
XX
PD   07-OCT-1999.
XX
PF   29-MAR-1999; 99WO-US007431.
XX
PR   30-MAR-1998; 98US-0079916P.
PR   16-OCT-1998; 98US-0104656P.
XX
PA   (NWBI-) NORTHWEST BIOTHERAPEUTICS INC.
XX

```
PI   Murphy GP, Boynton AL, Sehgal A;
XX
DR   WPI; 2000-052634/04.
DR   P-PSDB; AAY52507.
XX
PT   Use of CXCR-4 and SDF-1 as markers for diagnosis and treatment of e.g.
PT   tumors, degenerative disorders, growth deficiencies, hyper- and
PT   hypoproliferative disorders, physical trauma, lesions and wounds.
XX
PS   Disclosure; Fig 14; 138pp; English.
XX
CC   This sequence represents cDNA encoding CXC-chemokine receptor-4 (CXCR-4).
CC   Chemokine receptors play an important role in the chemotaxis of T-cells
CC   and phagocytic cells to areas of inflammation. CXCR-4 is a member of the
CC   G-protein-coupled receptor family, which is involved in signal
CC   transduction, and its ligand is stromal cell derived factor-1 (SDF-1,
CC   AAY52508). CXCR-4 also mediates CD4-independent infection by HIV-1. CXCR-
CC   4 has now been found to have a role in the aberrant proliferative
CC   behaviour of a number of cell types, including numerous primary tumours
CC   and derived cell lines. CXCR-4 is involved in cell transformation and
CC   tumorigenesis, particularly in brain, breast and colon tumours. It was
CC   found to be overexpressed in several brain tumour derived cell lines and
CC   primary brain tumour tissues, breast tumour tissues, colorectal
CC   adenocarcinoma, lung carcinoma and melanoma cell lines. CXCR-4 expression
CC   was required for the continuous proliferation of glioblastoma cancer
CC   cells, and inhibition of its gene function resulted in growth arrest.
CC   Conversely, overexpression resulted in enhanced and rapid cellular
CC   proliferation. CXCR-4 and SDF-1 can be used as markers for the diagnosis
CC   and prognosis of cell proliferative disorders, and antisense
CC   oligonucleotides complementary to at least a portion of an RNA transcript
CC   of a CXCR-4 gene can be used to inhibit hyperproliferation of a tumour
CC   cell. Agents that inhibit CXCR-4 function can be used for treating or
CC   preventing a disease or disorder involving cell overproliferation, e.g.,
CC   brain cancer, breast cancer, colon cancer, prostate cancer and B cell
CC   lymphoma, and also premalignant conditions, benign tumours,
CC   hyperproliferative disorders, and benign dysproliferative disorders. They
CC   can also be used for treating e.g., cirrhosis of the liver, keloid
CC   formation, psoriasis, benign tumors, fibrocystic conditions and tissue
CC   hypertrophy. Compounds that promote CXCR-4 function can also be used for
CC   preventing or treating a disease or disorder involving a deficiency in
CC   cell proliferation, or treating a condition where cell proliferation
CC   would be desirable. Such diseases include degenerative disorders (e.g.,
CC   Parkinson's disease, Alzheimer's disease), growth deficiencies,
CC   hypoproliferative disorders, physical trauma, lesions (e.g., those caused
CC   by ischaemia), and wounds
XX
SQ   Sequence 1059 BP; 244 A; 301 C; 232 G; 282 T; 0 U; 0 Other;

Query Match      75.0%;  Score 517.2;  DB 3;  Length 1059;
Best Local Similarity  85.8%;  Pred. No. 8.8e-137;
Matches  592;  Conservative  0;  Mismatches  83;  Indels  15;  Gaps  1;

Qy    1  CTGCACCTGTCAGTGGCTGGCTGACCTCCTCTTGTCATCACTCCCCTTCTGGGCAGTTGAT  60
         |||||||||||||||||||| |||||||||||||  |||||||| ||||||||||||||||
Db  232  CTGCACCTGTCAGTGGCTGGCCGACCTCCTCTTCTTTGTCATCACGCTTCCCCTTCTGGGCAGTTGAT  291

Qy   61  GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTCACACT  120
         ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  292  GCCGTGGCAAACTGGTACTTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTCACACA  351

Qy  121  GTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  352  GTCAACCTCTACAGCAGTCCTCATCCTCTCAGCTGGCCTTCATCAGTCTGGACCGGCTACCTCGGCC  411

Qy  181  ATTGTCACGCCACCAACAGTCAAAGCCAAGGAAAACTGCTGGCTGAAAAGGCAGCTCTAT  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  412  ATCGTCCAACGCCACCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT  471

Qy  241  GTGGGGCTCTGGATCCCAGCCCTCCTCCTGACTATACCTGGACTTCATCTTTGCCGACGTC  300
         ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  472  GTTGGGCGCTGGATCCCCTGCCCTCCTCCTGCTGACTATTCCCGACTATTCATCTTTGCCAACG--  529

Qy  301  AGCCAGGGGGACCATCAGTCAGTCAGGGGGATGACGAGATACATCTGTGACCGGCTTTACCCCGAT  360
```

```
Db  530  -----------TCAGTGAGGCAGCAGATGACGAGATATCTGTGACCGGCTTCTACCCCAAT  576
                     |||| |||   ||||  ||||||||||||| ||  |||||||||||||||||
Qy  361  AGCCTGTGATGGTGGTGTTCAATTCCAGCACATAAATGGTGGTCTCATCCTGCCCGGC  420
             |||| ||||||| |||||||||||| ||||||||| ||| ||||||||||||||||
Db  577  GACTGTGGGTGGTTGTGTTCCAGTTCAGCACATCATGGTTGGCCTATCCTGCCTGT  636

Qy  421  ATCGTCATCCTCCTCGTTACTGCCATCATCATCCTTCAAGCTGTCACACTCCAAGGGCCAC  480
         |||||||||||  ||| |||  |||||||||||||| ||||||||||||||||||||||||
Db  637  ATTGTCATCCTGCTCCTGCTATTGCATTGCATTATCATCCTCCAAGCTGTCACACTCCAAGGGCCAC  696

Qy  481  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTCATCCTTTCTTTTGCTCTGG  540
         |||||||||||||||||||||||  ||  |||||||||||||||||||| |||  ||||||
Db  697  CAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTCTTCCTTCTTCGCCTGTTGG  756

Qy  541  CTGCCATATATGTGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA  600
         ||||| |  |  | |||||||||||  |||||| |||   |||||| |||||||||||
Db  757  CTGCCTTACTACATTGGGATCAGCATCAGCATGACTCCTCCTGGAAATCATCAAGCAA  816

Qy  601  GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCCTCCATCACAGAGCCCTCGCCTTC  660
         || |||||| ||||| |  |||||||||| |||||| | ||||||| ||| ||| |||||
Db  817  GGGTGTGAGTTTGAGAACACTGTGCACAGTGGATTTCCATCACCGAGGGCCCTAGCTTTC  876

Qy  661  TTCCACTGTGCCTGAACCCCATCCTCTAT  690
         ||||||||| | ||||||||||||||||||
Db  877  TTCCACTGTTGTCTGAACCCCATCCTCTAT  906
```

```
RESULT 6
ABQ74938
ID   ABQ74938 standard; cDNA; 1059 BP.
XX
AC   ABQ74938;
XX
DT   30-OCT-2002 (first entry)
XX
DE   Human CXC chemokine receptor 4 encoding cDNA SEQ ID NO:125.
XX
KW   Cysteine-X-cysteine chemokine receptor 4; CXC chemokine receptor 4;
KW   binding; identification; AIDS; HIV; CXCR4; virucide; anti-HIV; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1059
FT                   /*tag= a
FT                   /product= "Cysteine-X-cysteine chemokine receptor 4"
XX
PN   WO200257313-A2.
XX
PD   25-JUL-2002.
XX
PF   26-OCT-2001; 2001WO-US051165.
XX
PR   27-OCT-2000; 2000US-0243587P.
PR   20-MAR-2001; 2001US-00813448.
PR   20-MAR-2001; 2001US-00813651.
PR   20-MAR-2001; 2001US-00813653.
XX
PA   (CONS-) CONSENSUS PHARM INC.
XX
PI   Nestor JJ, Wilson CJ, Tan Hehir CA, Kates SA, Kretenansky J;
XX
DR   WPI; 2002-608378/65.
DR   P-PSDB; ABP52651.
XX
PT   Identifying a binding compound for cysteine-X-cysteine chemokine receptor
PT   4 (CXCR4) for preventing or treating AIDS and HIV infection, comprises
PT   binding a molecule to a molecule with a binding property corresponding to
PT   CXCR4.
XX
PS   Example 3; Fig 3A-B; 92pp; English.
XX
```

RESULT 7

AAH99952

```
ID   AAH99952 standard; cDNA; 1059 BP.
XX
AC   AAH99952;
XX
DT   18-JUN-2002 (first entry)
XX
DE   CXCR4 encoding sequence.
XX
KW   Human Chemokine receptor 4; CXCR4; gene; human; ss; chromosome 2q21;
KW   HIV-1.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(414,T)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
FT   variation       replace(783,T)
FT                   /*tag= b
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200179229-A2.
XX
PD   25-OCT-2001.
XX
PF   13-APR-2001; 2001WO-US012268.
XX
PR   13-APR-2000; 2000US-0197025P.
XX
PA   (GENA-) GENAISSANCE PHARM INC.
XX
PI   Bentivegna SC,  Chew A,  Choi JY,  Koshy B;
XX
DR   WPI; 2002-075062/10.
DR   P-PSDB; AAG78471.
XX
PT   Isolated human chemokine receptor 4 polynucleotide, useful for studying
PT   expression and function of chemokine receptor 4 and for treating human
PT   immunodeficiency virus-1.
XX
PS   Claim 24; Fig 2; 54pp; English.
XX
CC   The invention relates to an isolated chemokine receptor 4 (CXCR4)
CC   polynucleotide which is useful for studying expression and function of
CC   chemokine receptor 4 and for the treatment of viruses, including HIV-1.
CC   Methods associated with chemokine receptor 4 are useful for improving the
CC   efficacy and reliability of several steps in the discovery and
CC   development of drugs for treating diseases associated with CXCR4
CC   activity, e.g., human immunodeficiency virus-1 infection. They can
CC   validate CXCR4 as a candidate agent for treating a specific condition or
CC   disease predicted to be associated with CXCR4 activity, and in the design
CC   of clinical trials of candidate drugs for treating a specific condition
CC   or disease predicted to be associated with CXCR4 activity. CXCR4
CC   polynucleotides are useful for therapeutic purposes such as to treat HIV-
CC   1 infection and also are useful for studying expression of the CXCR4
CC   isogenes in vivo. This sequence represents the CXCR4 encoding sequence.
XX
SQ   Sequence 1059 BP; 244 A; 301 C; 232 G; 282 T; 0 U; 0 Other;
```

Query Match   75.0%; Score 517.2; DB 6; Length 1059;
Best Local Similarity   85.8%;   Pred. No. 8.8e-137;
Matches   592; Conservative   0; Mismatches 83; Indels 15; Gaps 1;

```
Qy      1 CTGCACCTGTCAGTGGCTGACCTCCCTTCTGTCATCACACTCCCCTTCTCGGGCAGTTGAT   60
          ||||||||||||||||||||||||| ||||||||||||| || ||||||||||||||||||
Db    232 CTGCACCTGTCAGTGGCCGACCTCTCTGTCATCACGCTTCCCTTCCCTTCTGGGCAGTTGAT  291

Qy     61 GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT  120
          |||| ||||| ||||||||| |||||||  ||||| ||||||||||||| |||||||||||
Db    292 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  351

Qy    121 GTCAACCTCTACAGCAGCCGTCTCATCCTGGCCTTCAGCCTGGACGGGTACCTCGCC  180
          |||| ||||| ||||||||| |||||||||||||||||||||||| ||||||||||
Db    352 GTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTGTCTGGACCGCTACCTGGCC  411

Qy    181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAAACTGCTGCTGAAAAGCCAGTCTAT  240
          |||||||||||||||||||||||| |||||| |||| |||||||| ||||||||||||||
Db    412 ATCGTCCACGCCACCAACAGTCAGAGGCCCAGCAAGAAGCTGTTGCTGGTCGTTCTAT  471

Qy    241 GTGGGCGTCTGGATCCCAGCCCTCCTGCTATACCTGACTTCATCTTGCGGACGTC  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 GTTGGCGTCTGGATCCCTGCCCCTGCGACTTCATCTTTGCCAACG--  529

Qy    301 AGCCAGGGGGACATGCAGTCAGGGGATGACAGTACATCTGTGACCGCCTTACCCCGAT  360
          ||||| ||| ||||||||||||| |||||
Db    530 --------TCAGTGAGGCAGATGACAGATATATCTGTGACCGGCTTCTACCCCAAT  576

Qy    361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCAGATATAATGGTGGTCTCATCCTGCCCGGC  420
          || |||||| |||||||| ||||| ||||| ||||| |||||||||||||||||||||||
Db    577 GACTTGTGGGTGGTGTTGTTCCAGTTGCAGCACCATCATGGTTGGCCTTATCCTGCCTGGT  636

Qy    421 ATCGTCATCCTCTCGTTACTGGTCGCATCATCATCCTCTAAGCTGTCACACTCCAAGGGCAC  480
          || ||||||||| |||||| || ||||||||||||||||||||| |||||||||||||||||
Db    637 ATTGTCATCCTGTCGTTAGGATTGCATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCAC  696

Qy    481 CAGAAGCGCAAGGCCCTCAAGACAGCAGTGACACGTCATCCTCAGCAGTTCTTTGCCTGCTGG  540
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    697 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTCGCTTTCTTCGCCTGCTGG  756

Qy    541 CTGCCATATTATGTGGGGGATCAGCATCGGATCCTCCTTCATCCTTTGGGAGTCATCAAGCAA  600
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    757 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAA  816

Qy    601 GGATGTGACTTGCAGAGCATTGTGCACAAGTGACATCTGCACACTCACACGAGGCCCTGGCCTTC  660
          ||| |||||| |||||||||||||||| || |||||||||||||| ||||||||||| ||||||
Db    817 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC  876

Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
          |||||||||||||||||| |||||||||||
Db    877 TTCCACTGTTGTCTGAACCCCATCCTCTAT  906
```

CC The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
XX
SQ Sequence 1059 BP; 244 A; 300 C; 232 G; 283 T; 0 U; 0 Other;

Query Match   75.0%; Score 517.2; DB 7; Length 1059;
Best Local Similarity   85.8%; Pred. No. 8.8e-137;
Matches 592; Conservative   0; Mismatches 83; Indels 15; Gaps   1;

```
Qy    1  CTGCACCTGTCAGTGGCTGACCTCCCTCTTGTCATCACACTCCCCTTCTCTGGGCAGTTGAT   60
         ||||||||||||||||| |||||||||||| |||||||| || |||||||||||||||||||
Db  232  CTGCACCTGTCAGTGGCCGACCTCCCTCTTTGTCATCACGCTTCCCTTCTCTGGGCAGTTGAT  291

Qy   61  GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT  120
         ||| ||||  |||||||||||||||||  || || ||||||| || |||||||| ||||||
Db  292  GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGCAGTCATCATCTCACACA  351

Qy  121  GTCAACCTCTACAGCACGCGTTCTCATCCTGGCCTTCATCATTGCCGACGTC  180
         ||||||||||||||| |||  |||||||||||||||||||||  || |||
Db  352  GTCAACCTCTACAGCAGTCGCTCCTCATCCTGGCCTTCATCAGTGTGGCAACG--  411

Qy  181  ATTGTCCACGCCACCAACAGTCAAAGGCCAAGAAACTGCGTGCTGAAAAGGCAGTCCAT  240
         ||||| |||||||||||||||| | |||||||||||||| | | ||||||||||| ||||
Db  412  ATCGTTCACGCCACCAACAGTCAGAGGCCAAGAAGCTGTTGGCTGAAAAGGTGGTTCTAT  471

Qy  241  GTGGGCGTCTGGATCCCAGCCCTCCTCGTGACTATACCTGACTTCATCTTTGCCGACGTC  300
         || |||||||||||||| ||||| ||| | |||| |||  ||||||||||||||| |||
Db  472  GTTGGCGTCTGGATCCCGCCCCTGCCTGCTGACTTCCCGACTTCATCTTTGCCAACG-  529

Qy  301  AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT  360
         ||| | || | | | |  |||| ||| |||||| |||||||||||||||| |||||| ||
Db  530  --------TCAGTGAGGCAGATGACAGATATCTGTGACCGCCTTCTACCCCCAAT  576

Qy  361  AGCCTGTGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCCGC  420
         ||| |||||  |||||||| || || |||||| | |||| ||| || |||| |||| ||
Db  577  GACTTGTCGGGTGGTGTGTTCCAGTTTGCAGTTTCAGCATCATGTGGCCTTATCCTCCTCGT  636

Qy  421  ATCGTCATCCTCCTCCTGTTACTGCATCATCATCATCCTTAAGCTGTCACACTCCAAGGGCCAC  480
         || ||||||| ||| ||| || ||||| ||||| || |||| ||||||||||||||||||||||
Db  637  ATTGTCATCCTGTCCTGCTATTGCATTATCATCCCAAGGCTGTCACACTCCAAGGGCCAC  696

Qy  481  CAGAAGCGCAAGGCCCTCAAGACGCACAGTCATCCTCATCCTAGCTTTCTTGCTTGCTGG  540
         ||||||| ||||||||||||||| |||||||||||  ||||||  ||||||| || ||||
Db  697  CAGAAGCCAAGGCCCTCAAGACCACAGTCATCCCATCCTCCTGGCTTTCTTCTTCGGCCTGTTGG  756

Qy  541  CTGCCATATTATGTGGGGATCAGCATCGACATCGACATCCTTCCTCATCCTTTGGGAGTCATCAAGCAA  600
         ||| ||||| ||  ||||||||| | ||| |||| ||| |||||| |||||||||||| ||||||| |||
Db  757  GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGCACAAGTGGGATTTCACATCCTTCATCACCAGCTTTTC  816

Qy  601  GGATGTGACTTCGAGAGCAATTGTGCACAAGTGGGATCTCCATCACAGAGGCCCTCGCCTTC  660
         |||| | |||| |||||| || ||||||||||||||   ||||||||| |||||| | ||||
Db  817  GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGCACAAGTGGGATTTCCATCACACCGGGCCCTAGCTTTC  876

Qy  661  TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
         |||||||||| |||||||||||||||||||
Db  877  TTCCACTGTTGTCTGAACCCCATCCTCTAT  906
```

RESULT 8
ACC72707
ID   ACC72707 standard; cDNA; 1059 BP.
XX
AC   ACC72707;
XX
DT   09-JUL-2003 (first entry)
XX
DE   Human cancer related protein encoding cDNA SEQ ID NO:46.
XX
KW   Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW   heart disease; atherosclerosis; endometriosis; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003025138-A2.
XX
PD   27-MAR-2003.
XX
PF   17-SEP-2002; 2002WO-US029560.
XX
PR   17-SEP-2001; 2001US-0323469P.
PR   20-SEP-2001; 2001US-0323887P.
PR   13-NOV-2001; 2001US-0350666P.
PR   08-FEB-2002; 2002US-0355145P.
PR   08-FEB-2002; 2002US-0355257P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;

CC   identifications and in identifying promising antibiotics, antiviral or
CC   antifungal agents
XX
SQ   Sequence 1225 BP; 288 A; 324 C; 259 G; 354 T; 0 U; 0 Other;

Query Match   75.0%;   Score 517.2;   DB 6;   Length 1225;
Best Local Similarity 85.8%;   Pred. No. 9.3e-137;
Matches 592;   Conservative   0;   Mismatches 83;   Indels 15;   Gaps   1;

Qy   1 CTGCACCTGTCAGTGCTGACCTCCTTCTTTGTCATCACACTCCCCTTCTCGGCAGTTGAT 60
          ||||||||||||||||| ||||||||||| |||||||||| | |||||||||||||||||
Db   257 CTGCACCTGTCAGTGCCGGACCTCCCTCCTTGTCATCACGCTCCCTTCTCGGCAGTTGAT 316

Qy   61 GCCATGGCTGCACTGGTACTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACT 120
          ||| |||| |||||||||| |||||| || || ||| |||||| | ||||| ||||||
Db   317 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCATGTCATCTACACA 376

Qy   121 GTCAACCTTCACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
          ||||||||||||||||| || |||||||||||||||||||||| |||||||||||||||||
Db   377 GTCAACCTTCACAGACAGTGTCCTCATCCTGGCCTTCATCAGTCTGACCGGTACCTCGGCC 436

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
          || ||||| |||||||||||||| ||||||||| |||| ||||||||||| | |||||||
Db   437 ATCGTCCAGGCCACCAACAGTCAGAGGCCAAGGAGGAGCTGTTGGCTGAAAAGGTGGTCTAT 496

Qy   241 GTGGGGCTGTCGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
          || ||||||| ||||| ||| ||||| ||||||||| ||||||||||| ||||||||
Db   497 GTTGGGGCTGGATCCCTGCCGTCCTCGCTGCTGGACTATTCCCGACTTCATCTTTGCCAACG-- 554

Qy   301 AGCCAGGGGACATCAGTCAGGGGATGAGCAGGTACATCTGTGACCGCCTTTACCCCGAT 360
                      |||||| |||| ||| ||| | ||| ||||||||||||| ||||||||
Db   555 ------------TCAGTGAGCGCAGATGACAGATATATCTGTGACCGCCTTCTACCCCCAAT 601

Qy   361 AGCCTGCGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC 420
          ||| ||| | |||||| ||||| |||||||||| || | |||| |||||||||||||||
Db   602 GACTTGTGGGGTGGTTGTGTTTCAGTTTCAGCACATCATGTCGTTGGCCTTATCCTGCCTGGT 661

Qy   421 ATCGTCATCCTCTCGTTACTGCATCATCATCTCCTAAGCTGTCACACTCCAAGGGCCAA 480
          || ||||||||| || || ||||||||||||||| ||||||||||||||||||||||||
Db   662 ATTGTCATCCTCGTCCGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 721

Qy   481 CAGAAGCGCAAGGCCCTCAAGACAGCAGTCATCCTCATCCTCATCCTTGCCTGCTGG 540
          |||||||||||||||||||||||||| | |||||||||||||||||||||| | |||
Db   722 CAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTCATCCTGGCTTTCTCGCCTGTTGG 781

Qy   541 CTGCCATATTATGTGGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAA 600
          ||||| || | | |||||||||||||||||||||| |||| |||||||||||
Db   782 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAA 841

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
          || ||||| || |||| ||| ||||||||||||||| |||||||| ||||||||| | |||
Db   842 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 901

Qy   661 TTCCACTGTGCCTGAACCCCATCCTCTAT 690
          ||||||||||| |||||||||||||||||
Db   902 TTCCACTGTGTCTGAACCCCATCCTCTAT 931

RESULT 10
ACA56637
ID   ACA56637 standard; cDNA; 1225 BP.
XX
AC   ACA56637;
XX
DT   06-JUN-2003 (first entry)
XX
DE   Human signalling pathway polynucleotide probe SEQ ID NO 1235.
XX
KW   Human; probe; ss; array element; Parkinson's disease;
KW   signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW   immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS   Homo sapiens.

Qy   541 CTGCCATATTATGTGGGGATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA 600
          ||||| || | | ||||||||||||||||||||||| ||||| |||| |||||||||||
Db   757 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAA 816

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
          || ||||| || |||| ||| ||||||||||||||| |||||||| ||||||||| | ||
Db   817 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 876

Qy   661 TTCCACTGTGCCTGAACCCCATCCTCTAT 690
          ||||||||||| |||||||||||||||||
Db   877 TTCCACTGTGTCTGAACCCCATCCTCTAT 906

RESULT 9
ABZ35630
ID   ABZ35630 standard; cDNA; 1225 BP.
XX
AC   ABZ35630;
XX
DT   05-FEB-2003 (first entry)
XX
DE   Human gene expression profile polynucleotide SEQ ID NO 741.
XX
KW   Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW   bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW   tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW   gene expression; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200274979-A2.
XX
PD   26-SEP-2002.
XX
PF   20-MAR-2002; 2002WO-US008456.
XX
PR   20-MAR-2001; 2001US-0276947P.
XX
PA   (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI   Wan J, Wang Y;
XX
DR   WPI; 2002-740862/80.
XX
PT   New gene expression profile generated from primary, endothelial,
PT   epithelial, and muscle cell types, useful for identifying disease
PT   pathologies involving alterations of gene expression, e.g. cancer.
XX
PS   Example 3; Page 827-828; 850pp; English.
XX
CC   The invention relates to a gene expression profile comprising one or more
CC   genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
CC   is a coronary artery endothelium, umbilical artery or vein endothelium,
CC   aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC   endothelium, myometrium microvascular endothelium, keratinocyte
CC   epithelium, bronchial epithelium, mammary epithelium, prostate
CC   epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC   small airway epithelium, renal epithelium, umbilical artery smooth
CC   muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC   dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC   aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC   bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC   osteoblasts or prostate stromal cell. The gene expression profile is used
CC   for determining the level of RNA expression for a sample, determining the
CC   phenotype of a cell and distinguishing cell types. The gene or a protein
CC   expression profile is useful in identifying disease pathologies involving
CC   alterations of gene expression. The assessment of expression profiles may
CC   provide meaningful information with respect to tumour type and stage,
CC   treatment methods, and prognosis. The gene or protein expression profile
CC   may also be used for creating microarrays. The microarray is useful for
CC   genetic and physical mapping of genomes, DNA sequencing, genetic or
CC   medical diagnosis, genotyping of organisms, confirming cell or tissue

```
XX
PN   US6500938-B1.
XX
PD   31-DEC-2002.
XX
PF   30-JAN-1998;   98US-00016434.
XX
PR   30-JAN-1998;   98US-00016434.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Au-Young J,  Seilhamer JJ;
XX
DR   WPI; 2003-352189/33.
XX
PT   Combination of polynucleotide probes, useful as array elements in a
PT   microarray for monitoring the expression of a number of target
PT   polynucleotides.
XX
PS   Claim 1; SEQ ID NO 1235; 65pp; English.
XX
CC   The invention relates to a combination which, comprises a number of
CC   polynucleotide probes comprising a sequence selected from one of the 1490
CC   sequences mentioned in the specification. The combination is useful as an
CC   array element in a microarray for monitoring the expression of a number
CC   of target polynucleotides. The microarray is particularly useful in the
CC   diagnosis and treatment of cancer and immunopathology and neuropathology.
CC   The microarray is useful in diagnostics and treatment regimens, drug
CC   discovery and development, toxicological and carcinogenicity studies,
CC   forensics and pharmacogenomics. the microarray is also useful for
CC   monitoring progression of diseases and for developing sophisticated
CC   profiles for the effects of currently available therapeutic drugs. The
CC   combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC   and genomic fragments and in research and diagnostic applications. The
CC   array can detect changes in expression in a large number of genes coding
CC   for different signaling pathway populations which can be used to diagnose
CC   various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC   immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC   and Parkinson's disease. The present sequence represents a polynucleotide
CC   probe of the invention. Note:The sequence data for this patent did not
CC   form part of the printed specification but was obtained in electronic
CC   format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=06500938B1
XX
SQ   Sequence 1225 BP; 288 A; 324 C; 259 G; 354 T; 0 U; 0 Other;

Query Match             75.0%;  Score 517.2;  DB 7;  Length 1225;
Best Local Similarity   85.8%;  Pred. No. 9.3e-137;
Matches  592;  Conservative    0;  Mismatches   83;  Indels   15;  Gaps    1;

Qy     1 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCCGGGCAGTTGAT 60
           ||||||||||||||||||  |||||||||| ||||||| |||||||  ||||||||||||
Db   257 CTGCACCTGTCAGTGGCTGGCCGACCTCCTTCCTTCCTTGTCATCACGCTTCCCTTCCGGGCAGTTGAT 316

Qy    61 GCCATGGCTGACTGTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTCACACT 120
           || ||||| |||||||||||||||||    ||||| ||||||||||||| |||||| |||||
Db   317 GCCGTGGCAAACTGGTACTTTGGGACTTCCTATGCAAGGCAGTCCATGTCATCTCATCTACACA 376

Qy   121 GTCAACCTCTACAGCAGGAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCCGCC 180
           |||||||||||||||| |||  | || |||||||||||||||| ||||||||||||||||||||
Db   377 GTCAACCTCTACAGCAGTGCCCTATCCCGGCCTTCATCAGTCTGGACCGGTACCTCCGCC 436

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
           || |||||||||||||||||||| ||||| |||||| ||| |||||||||||| |||||||
Db   437 ATCGTCCACGCCACCAACAGTCAGAGGCGGCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT 496

Qy   241 GTGGGGCGTCTGGATCCCAGCCCTCCTCCGACTATACCTGACTTCATCTTTGCCGACGTC 300
           || |||| ||||||||| |||||||||  || || ||| ||| ||||||||||||| |||
Db   497 GTTGGGGTCTGGATCCCCGCCCTCCTGCCCTGCTGACTATTCCCGACTTCATCTTTGCCAACG-- 554

Qy   301 AGCCAGGGGGACACATGCAGGGGATGACAGGGGATACATCGTGACCGCCTTTACCCCGAT 360
                     |||||| ||||||| |||||| ||| |||||||||||||||||| |||||| ||
Db   555 ----------TCAGTGAGGCAGATGACGATGACAGATATATCGTGACCGCCTTCTACCCCCAAT 601

Qy   361 AGCCTGTGGATGGTGGTGTTCAATTCAGCAGCATATATAATGTGGGTCTCATCGTGCCCGC 420
           | |||||||| |||| ||||| |||| ||||||||| |||| ||| ||| |||| ||||||
Db   602 GACTTGTGGGTGGGTGGTTGTTGTTCCAGTTTCAGCACATCATGATGTTGGCCTTATCCTGCCTGGT 661

Qy   421 ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCAAGCTGTCAAGGGGCCAC 480
           || |||||||| ||||| |||| |||||||||||| |||||||| || ||||||||| ||
Db   662 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGGCAC 721

Qy   481 CAGAAGCGGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTCAGCTTTCTTTGCCTCTGG 540
           ||||||||||||||||||||||| || |||||||||||||||||  |||  | |||| |||
Db   722 CAGAAGCGGCAAGGCCCTCAAGACCCACAGTCATCCTCATCCTCCTTTTGGGAGTCATCAAGCAA 781

Qy   541 CTGCCATATTATGTGGGATCAGCATCAGCATGACTCCTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
           ||||| || |||  |||||||||||||||| |||||||||||||||||  | ||||||||||||||
Db   782 CTGCCTTACTACATTGGATCAGCATCAGCATCGACTCCTCCTTCATCCTCCTGGAAATCATCAAGCAA 841

Qy   601 GGATGTGGACTTCGAGGACCATTGTGCCACAAGTGGATCTCCATCACAGAGGCCCGCCCTC 660
           || |||| | |||||| || | ||||| |||||| |||| ||||||| |||||||   |||
Db   842 GGGTGTGAGTTTGAGAACACCTGTGCACAAGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 901

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
           |||||||||||  ||||||||||||||||||
Db   902 TTCCACTGTTGTCTGAACCCCATCCTCTAT 931

RESULT 11
AAV18357
ID   AAV18357 standard; DNA; 1317 BP.
XX
AC   AAV18357;
XX
DT   25-SEP-1998  (first entry)
XX
DE   Human RM3 seven transmembrane (7TM) receptor cDNA.
XX
KW   V28; placenta; seven transmembrane receptor; 7TM; signal transduction;
KW   immunology; inflammation; RM3; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             201..1211
FT                   /*tag= a
FT                   /product= "Human RM3 seven transmembrane receptor"
XX
PN   US5759804-A.
XX
PD   02-JUN-1998.
XX
PF   17-NOV-1993;   93US-00153848.
XX
PR   17-NOV-1992;   92US-00977452.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Schweickart VL,  Godiska R,   Gray PW;
XX
DR   WPI; 1998-332132/29.
DR   P-PSDB; AAW48734.
XX
PT   DNA encoding V28 seven transmembrane receptor polypeptide - useful for
PT   producing recombinant polypeptide and anti-V28 antibodies, and in
PT   screening assays for V28 agonists and antagonists.
XX
PS   Example 11; Col 89-94; 56pp; English.
XX
CC   The present novel sequence represents the human RM3 cDNA isolated from a
CC   human macrophage cDNA library. The invention claims for the full length
CC   V28 genomic DNA sequence (AAV18343) isolated from a human placenta
CC   genomic library. The V28 (AAW48722) and RM3 proteins are seven
CC   transmembrane (7TM) receptors which are probably involved in signal
CC   transduction. The invention also claims that cells transformed with V28
```

Left column:

```
CC   DNA can be used to produce the recombinant polypeptide, to produce anti-
CC   V28 antibodies or in screening assays for V28 agonists or antagonists.
CC   The antibodies, agonists and antagonists could then be used to modulate
CC   V28 receptor-ligand binding, for e.g. in immunological and/or
CC   inflammatory events in vivo
XX
SQ   Sequence 1317 BP; 332 A; 342 C; 265 G; 378 T; 0 U; 0 Other;

Query Match             75.0%;  Score 517.2;  DB 2;  Length 1317;
Best Local Similarity   85.8%;  Pred. No. 9.6e-137;
Matches   592;  Conservative   0;  Mismatches   83;  Indels  15;  Gaps   1;

Qy      1  CTGCACCTGTCAGTGGCTGACTCCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
           |||||||||||||||| ||||| |||||||||||||||| || ||||||||||||||
Db    387  CTGCACCTGTCAGTGGCCGAACTCCTCTTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  446

Qy     61  GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
           |||||||||||| |||| ||||||||| |||||||||||||||| ||||||||||||||
Db    447  GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  506

Qy    121  GTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCGGTACCTCGCC  180
           |||||||||||||||||  |||||||| ||||||||||||| |||||| |||
Db    507  GTCAACCTCTACAGCAGTGCCTCATCAGTGGCCTTCATCAGTCTGGACCTCGGCC  566

Qy    181  ATTGTCCACGCCACCAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
           |||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db    567  ATCGTCCACGCCACCAGTCAAGGCCAACGGCCAAGGAAAGTGGTCTAT  626

Qy    241  GTGGGCCTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
           |||||| |||||||||||| ||||| ||||||||||||||||||| ||||||||||||
Db    627  GTTGGCCGTCTGGATCCCGCCCCCTGTGACTTCATCTTTGCCACG--  684

Qy    301  AGCCAGGGGGACTCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTCTACCCCGAT  360
           ||||||                ||||| |||||||||||||||||||||||| ||||
Db    685  ----------TCAGTGAGGCAGATGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT  731

Qy    361  AGCCTGTGGATGGTGGTGTTCAATTCCAGCACATATAATGGTGGTCTCATCCTGCCCGGC  420
           |||||| ||| ||||| |||||||||||| ||||||||||||||||||||||| ||| |
Db    732  GACTTGTGGGTGGTTGTGTTCCAGTTTCAGCACACATCTTCAGCACCACATCATCTGCCTGGT  791

Qy    421  ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGCCATCAT... 480
Db    792  ATTGTCATCCGTCTGCTATTCGCATCATCATCTCTTCGCATCAT... 851

Qy    481  CAGAAGGCGAAGGCCCTCAAGACGACGCAGTCATCATCCTTCTTTGCCTGCTGG  540
Db    852  CAGAAGGCGAAGGCCCTCAAGACCAACGACAGTGGCCTCTTTCTTGCCTTGTGG  911

Qy    541  CTGCCATTATGTGGGATCAGCATCGCTCCTTTTGGGAGTCATCAAGCAA  600
Db    912  CTGCCTTACTACATTGGGATCATCAGCATCAGCATCCTGGAAATCAAGCAA  971

Qy    601  GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACACAGAGGCCCTGCCTTC  660
Db    972  GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGGAGCCCTAGCTTTC  1031

Qy    661  TTCCACTGTTGCCTGAACCCATCCTCTAT  690
Db   1032  TTCCACTGTTGTCTGAACCCCATCCTCTAT  1061

RESULT 12
AAA91726
ID   AAA91726 standard; cDNA; 1317 BP.
XX
AC   AAA91726;
XX
DT   26-JAN-2001  (first entry)
XX
DE   Human 7TM receptor RM3 cDNA clone.
XX
KW   Seven transmembrane receptor; 7TM; heptahelical; serpentine;
KW   G-protein-coupled; V28; V31; V12; R20; R2; R12; RM3; gene therapy;
```

Right column:

```
KW   cancer; ss.
XX
OS   Homo sapiens.
PN   US6107475-A.
PD   22-AUG-2000.
XX
PF   26-APR-1999;   99US-00299843.
XX
PR   17-NOV-1992;   92US-00977452.
PR   17-NOV-1993;   93US-00153848.
PR   17-MAY-1994;   94US-00245242.
PR   01-JUN-1998;   98US-00088337.
PA   (ICOS-) ICOS CORP.
XX
PI   Schweickart VL,  Gray PW,  Godiska R;
XX
DR   WPI; 2000-571335/53.
DR   P-PSDB; AAB21698.
XX
PT   Polynucleotide encoding seven transmembrane receptors, antibody specific
PT   to the receptor, agonist and antagonist of the receptor useful for
PT   treating inflammation in a mammal.
XX
PS   Example 11; Col 95-98; 61pp; English.
XX
CC   The present sequence is a novel coding sequence for a seven transmembrane
CC   (7TM) receptors (also known as heptahelical, serpentine or G-protein-
CC   coupled receptors). The present sequence may be used for gene therapy for
CC   diseases such as cancer
XX
SQ   Sequence 1317 BP; 332 A; 342 C; 265 G; 378 T; 0 U; 0 Other;

Query Match             75.0%;  Score 517.2;  DB 3;  Length 1317;
Best Local Similarity   85.8%;  Pred. No. 9.6e-137;
Matches   592;  Conservative   0;  Mismatches   83;  Indels  15;  Gaps   1;

Qy      1  CTGCACCTGTCAGTGGCTGACTCCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
           |||||||||||||||| ||||| |||||||||||||||| || ||||||||||||||
Db    387  CTGCACCTGTCAGTGGCCGAACTCCTCTTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  446

Qy     61  GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
           |||||||||||| |||| ||||||||| |||||||||||||||| ||||||||||||||
Db    447  GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  506

Qy    121  GTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCGGTACCTCGCC  180
           |||||||||||||||||  |||||||| ||||||||||||| |||||| |||
Db    507  GTCAACCTCTACAGCAGTGCCTCATCAGTGGCCTTCATCAGTCTGGACCTCGGCC  566

Qy    181  ATTGTCCACGCCACCAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
           |||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db    567  ATCGTCCACGCCACCAGTCAAGGCCAACGGCCAAGGAAAGTGGTCTAT  626

Qy    241  GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
           |||||| |||||||||||| ||||| ||||||||||||||||||| ||||||||||||
Db    627  GTTGGCCGTCTGGATCCCGCCCCCTGTGACTTCATCTTTGCCACG--  684

Qy    301  AGCCAGGGGGACTCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTCTACCCCGAT  360
           ||||||                ||||| |||||||||||||||||||||||| ||||
Db    685  ----------TCAGTGAGGCAGATGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT  731

Qy    361  AGCCTGTGGATGGTGGTGTTCAATTCCAGCACATATAATGGTGGTCTCATCCTGCCCGGC  420
           |||||| ||| ||||| |||||||||||| ||||||||||||||||||||||| ||| |
Db    732  GACTTGTGGGTGGTTGTGTTCCAGTTTCAGCACACATCTTCAGCACCACATCATCTGCCTGGT  791

Qy    421  ATCGTCATCCTCCTGTTACTGCATCATCATCTCTTAAGCTGCCATCAT... 480
Db    792  ATTGTCATCCGTCTGCTATTCGCATCATCATCTCTTCAGCATCAT... 851

Qy    481  CAGAAGGCGAAGGCCCTCAAGACGACGCAGTCATCATCCTTCTTTGCCTGCTGG  540
Db    852  CAGAAGGCGAAGGCCCTCAAGACCAACGACAGTGGCCTCTTTCTTGCCTTGTGG  911
```

```
CC    polynucleotide is useful in DNA/RNA hybridisation assays to detect the
CC    capacity of cells to synthesise a 7TM receptor. The host cell is useful
CC    for large scale production of 7TM receptors, and in assays for
CC    identifying antagonist or agonists of 7TM receptor binding. Other 7TM
CC    receptors isolated in this study include V31 (located on chromosome 17q12
CC    -21.2, associated with familial brest cancer), V28 (Chromosome 3p21-ter,
CC    associated with von Hippel-Lindau syndrome, thyroid hormone resistance,
CC    small cell cancer of the lung and pseudo-Zellweger syndrome), R20
CC    (chromosome 11 p11-p13 associated with hypoprothrombinaemia and
CC    dysprothrombinaemia), V112 and R12. The present sequence is a
CC    polynucleotide encoding a 7TM receptor of the invention
XX
SQ    Sequence 1317 BP; 332 A; 342 C; 265 G; 378 T; 0 U; 0 Other;

Query Match             75.0%;  Score 517.2;  DB 6;  Length 1317;
Best Local Similarity   85.8%;  Pred. No. 9.6e-137;
Matches  592;  Conservative   0;  Mismatches   83;  Indels   15;  Gaps    1;

Qy      1 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACCTCCCTTCTCGGGCAGTTGAT    60
          ||||||||||||||||||||||||||||| || |||||||||||| |||||||||||
Db    387 CTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCACGCTTCCCTTCTCGGGCAGTTGAT   446

Qy     61 GCCATGGTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT    120
          ||||||||||||||||| || |||||||| ||||||||||||||||||||||||||||
Db    447 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA   506

Qy    121 GTCAACCTCTACAGCACGCGTTCCATCCTGGCTTCATCAGCCTGGACCGGTACCTCGCC    180
          ||||||||||||||||| ||||||||||||| |||||||||||||||||||||||||||
Db    507 GTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGCTACCTGCC   566

Qy    181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGAAAAGGCAGTCTAT    240
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    567 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGCTGAAAAGGTGGTCTAT   626

Qy    241 GTGGGCGTCTGGATCCCAGCCCTTCCTCCTGCTATACCTGACTTCATCTTTGCCGACGTC    300
          ||||||||||||||||||| |||||||||| |||||||||||||| ||||||||||
Db    627 GTTGGGGTCTGGATCCCTGCCCTTCCTGCTATTCCCGACTTCATCTTTGCCAACG--   684

Qy    301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGCAGTGACATCTGTGACCGGCCTTACCGGAT    360
          |||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db    685 --------TCAGTGAGGCAGATGACAGATATATCTGTGACCGGCTTCTACCCCAAT   731

Qy    361 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGTCTCATCCTGCGCCGGC    420
          ||||||||||||||||||| ||||||| |||||| ||||| |||| || ||||||||
Db    732 GACTTGTGGGGTGGTTGTTGTTTCAGTTTCAGCACATCATGTCCTTATCCTGCCTGT   791

Qy    421 ATCGTCATCCTCCTCCTGTTACTGCATCATCATCCTCAAGCTGTCACACTCCAAGGGGCCAC    480
          |||||||||||||||| ||||| |||||| ||||||| |||||||||||||||||||
Db    792 ATTGTCATCCTGTCCTGCTATTGCATTATCATCCTCCAAGCTGTCACACTCCAAGGGGCCAC   851

Qy    481 CAGAAGGCCAAGGCCCTCAAGACGCCTCCTCATCCTCATCCTCATCCTTTCTTTGCTGCTGG    540
          ||||||||||||||||||||||||| |||||||||||||||||||||| |||||||||
Db    852 CAGAAGGCCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTCGGCTTTCTCGCCTGTTGG   911

Qy    541 CTGCCATATTATGTGGGGATCAGCATCGACTCCTTTTGGGAGTCATCAGCAA    600
          ||||||||||||||||| |||||||||||| |||||||||||| ||||||||
Db    912 CTGCCTTACTACATTGGGGATCATCGACATGACTCCTGGAAATCATCAAGCAA   971

Qy    601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCCTCCATCACAGAGGCCCTCGCCTTC    660
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    972 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCAGAGGCCCTAGCTTTC   1031

Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT    690
          ||||||||||||||||||||||||||||
Db   1032 TTCCACTGTTGCTGAACCCCATCCTCTAT   1061

RESULT 14
ABS64793
ID    ABS64793  standard; DNA; 1582 BP.
XX
AC    ABS64793;
XX
```

```
Db    852 CAGAAGGCCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTCGGCTTTCTCGCCTGTTGG   911
          ||||||||||||||||| |||||||
Qy    541 CTGCCATATTATGTGGGGATCAGCATCGACTCCTTTTGGGAGTCATCAGCAA    600
          ||||| ||| || |||| |||||||||||| || |||||||||| |||||||||
Db    912 CTGCCTTACTACATTGGGGATCATCGACATGACTCCTGGAAATCATCAAGCAA    971
Qy    601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCCTCCATCACAGAGGCCCTCGCCTTC    660
          ||||||||||| |||||| |||||||||||||||||| ||||||||| |||||||| ||||
Db    972 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCAGAGGCCCTAGCTTTC   1031
Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT    690
          |||||||||| ||||||||||||||||||
Db   1032 TTCCACTGTTGCTGAACCCCATCCTCTAT   1061

RESULT 13
ABK54255
ID    ABK54255  standard; cDNA; 1317 BP.
XX
AC    ABK54255;
XX
DT    18-JUN-2002  (first entry)
XX
DE    Human 7 transmembrane domain receptor RM3 cDNA.
XX
KW    Human; ss; gene; 7 transmembrane domain receptor; 7TM; antiinflammatory;
KW    immunomodulatory; immunological disease; inflammation;
KW    familial brest cancer; von Hippel-Lindau syndrome;
KW    thyroid hormone resistance; small cell cancer of the lung;
KW    pseudo-Zellweger syndrome; hypoprothrombinaemia; dysprothrombinaemia.
XX
OS    Homo sapiens.
XX
PN    US6348574-B1.
XX
PD    19-FEB-2002.
XX
PF    01-JUN-1998;  98US-00088337.
XX
PR    17-NOV-1992;  92US-00977452.
PR    17-NOV-1993;  93US-00153848.
PR    17-MAY-1994;  94US-00245242.
XX
PA    (ICOS-) ICOS CORP.
XX
PI    Godiska R,  Gray PW,  Schweickart VL;
XX
DR    WPI; 2002-266543/31.
DR    P-PSDB; AAU91240.
XX
PT    New seven transmembrane receptor polypeptide R2, useful for producing
PT    antibodies capable of modulating ligand/receptor binding reactions
PT    involved in in vivo immunological and/or inflammatory events.
XX
PS    Example 11; Col 95-98; 61pp; English.
XX
CC    The invention relates to a purified and isolated R2 seven transmembrane
CC    (7TM) receptor. Also included are an anti-R2 antibody, a hybridoma
CC    producing the antibody, antigenic peptide fragments of R2, an antiserum
CC    produced by immunising a mammal with a composition comprising R2 or its
CC    fragment, where the fragment comprises at least one R2 extracellular or
CC    intracellular domain, and obtaining antiserum from the mammal after the
CC    immunising step, where the antiserum contains antibodies that bind to the
CC    R2 7TM receptor, a polynucleotide encoding 7TM receptor R2 and a host
CC    cell stably transformed or transfected with the polynucleotide allowing
CC    the expression of R2 in the host cell. The antibody is useful for
CC    modulating ligand/antiligand binding of a R2 7TM receptor, especially
CC    those ligand/receptor binding reactions involved in immunological and/or
CC    inflammatory events in vivo. R2 is useful for producing antibodies. The
CC    antibody is useful for immunisation to generate anti-idiotypic
CC    antibodies, for purifying R2 polypeptides, for identifying cells
CC    producing the polypeptides on their surfaces and for the detection and
CC    quantification of 7TM receptors on cell surfaces and in fluids. The
```

```
DT   15-NOV-2002 (first entry)
XX
DE   CXCR4 receptor DNA sequence.
XX
KW   Receptor; virus infection; immune response; autoimmune disease;
KW   TNFR1 receptor; inflammation; autoimmune encephalomyelitis;
KW   rheumatoid arthritis; myocarditis; CXCR4 receptor; ds.
XX
OS   Unidentified.
XX
PN   WO200257308-A2.
XX
PD   25-JUL-2002.
XX
PF   22-JAN-2002; 2002WO-GB000246.
XX
PR   22-JAN-2001; 2001GB-00001576.
PR   07-FEB-2001; 2001GB-00003032.
XX
PA   (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI   Moore M, Isalan M, Reynolds L, Ullman C, Girdlestone J;
PI   Demaison C, Choo Y;
XX
DR   WPI; 2002-590720/63.
XX
PT   New polypeptide capable of binding nucleic acids for treating or
PT   preventing a disease caused by a virus or a disease associated with an
PT   immune response, particularly autoimmune diseases e.g. inflammation or
PT   rheumatoid arthritis.
XX
PS   Disclosure; Page 12-13; 105pp; English.
XX
CC   The present invention relates to a new polypeptide capable of binding to
CC   a nucleic acid comprising a receptor nucleotide sequence which can
CC   function as a receptor for virus infection or is involved in an immune
CC   response. The molecules of the invention are useful in the treatment or
CC   prevention of a disease caused by a virus or of a disease associated with
CC   an immune response, particularly an autoimmune disease. The zinc finger
CC   polypeptide or a nucleic acid encoding such a polypeptide can be used to
CC   modulate transcription of a receptor nucleotide sequence. The nucleic
CC   acid polypeptide capable of binding to a nucleic acid sequence comprising
CC   a receptor nucleotide sequence, where the receptor is capable of
CC   functioning as a receptor for infection by the virus or is involved in an
CC   immune response, or a nucleic acid encoding such a polypeptide, can be
CC   used in the preparation of a medicament for use in the treatment or
CC   prevention of a disease caused by a virus or of a disease associated with
CC   an immune response. The polypeptides may further be used to treat or
CC   prevent various diseases or syndromes associated with or caused by
CC   malfunction of the receptor as the TNFR1 receptor, such as autoimmune
CC   diseases including inflammation, autoimmune encephalomyelitis, rheumatoid
CC   arthritis or myocarditis. The present nucleic acid sequence represents a
CC   CXCR4 receptor sequence, as described in the invention
XX
SQ   Sequence 1582 BP; 386 A; 376 C; 333 G; 487 T; 0 U; 0 Other;

Query Match             75.0%;  Score 517.2;  DB 6;  Length 1582;
Best Local Similarity   85.8%;  Pred. No.1e-136;
Matches  592; Conservative    0; Mismatches   83; Indels   15; Gaps    1;

Qy     1 CTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
           ||||||| |||||||||||||| || || ||||||||| |||||||||||||||||||||
Db   223 CTGCACCTGTCAGTGGCCGACCCTCCTCCTTGTCATCACGCCTTCCCTTCTGGGCAGTTGAT 282

Qy    61 GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT 120
           |||||| ||||||  |||||| ||||||| ||  |||||||||||||||||||||| |
Db   283 GCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGAGTCCATGTCATCATCACACA 342

Qy   121 GTCAACCTCTACAGCAGGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
           |||||||||| ||||| || |||| |||||||||||||||||||||||||||||||||||
Db   343 GTCAACCTCTACAGCAGTGCCCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 402

Qy   181 ATTGTCCACGCACCACCAGTCAAAGGCCAAGGAAACTGCTGGCTAGAAAAGGCAGTCTAT 240
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   403 ATCGTCCAGGCCACCACAGTCAGAGGCCAAGGAAGCTGTGGCTGACTTCATCTTTGCCGACGTC 462

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
           || |||||||||||||| || ||||| |||||| || ||||||||||||||||  --
Db   463 GTTGGGCGTCTGGATCCCGCTCCTGCTGACTATTCCCGACTTCATCTTTGCCAACG-- 520

Qy   301 AGCCAGGGGACATCAGTCAGGGGATGACAGGGTACATCGTGACCGCCTTACCCGAT 360
           -----------||||| ||| ||||||||||| |||||||||||| || ||||||
Db   521 -----------TCAGTGAGGCAGATGACAGATATATCGTGACCGGCTTCTACCCCAAT 567

Qy   361 AGCCTGTGGATGGTGGTGTTCAATTCAGCAGCATATAATGGTGGTTCTCATCCTGCCCGGC 420
           |  |||||| ||||| |||||| || |||||||||| | ||||| |||||||||||||
Db   568 GACTTGTGGGTGGTTGTGTTCCAGTTTCCAGCACACATCATGGTTGGCCTTATCCTGCCTGGT 627

Qy   421 ATCGTCATCCTCTCTGTTACTGTCATCATCATCTCTCTAAGCTGTCACACTCCAAGGGGCCAC 480
           || ||||||||  || ||| |  || ||||||||||  ||||||||||||||| |||||||||
Db   628 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCCAAGGGGCCAC 687

Qy   481 CAGAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGG 540
           ||||| ||||||||||||||||  || |||||||||||||| |||||||  ||||| |||
Db   688 CAGAGGCGCAAGGCCCTCAAGACCACACAGTCATCCTCATCCTGGCTTTCTTCTTCGCCTGTTGG 747

Qy   541 CTGCCATATATGTGGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
           ||||| || || |||||||||||||||||| |||||||||||  |||| ||||||||||
Db   748 CTGCCTTACTACATTGGGGATCAGCATCGACATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAA 807

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAGTGGATCTCCATCACAGAGGCCTCGCCTTC 660
           ||  ||||  || ||||  || |||||| |||| |  ||||||| |||||||  ||||
Db   808 GGGTGGAGTTTGAGAACACTGTGCACAAGTGGAGTTTTCCATCACACCGGAGGCCCTAGCTTTC 867

Qy   661 TTCCACTGTTGCTGGAACCCCATCCTCTAT 690
           ||| ||||||||||||||||||||||||||
Db   868 TTCCACTGTTGCTGGAACCCCATCCTCTAT 897


RESULT 15
AA227610
ID   AA227610 standard; DNA; 1588 BP.
XX
AC   AA227610;
XX
DT   16-DEC-1999 (first entry)
XX
DE   Human CXCR4 coding sequence.
XX
KW   CXCR4; human; neovascularisation; inhibitor; solid cancer; therapy;
KW   tissue repairing agent; vascularisation; ss.
XX
OS   Homo sapiens.
XX
PN   WO9948528-A1.
XX
PD   30-SEP-1999.
XX
PF   23-MAR-1999;  99WO-JP001448.
XX
PR   24-MAR-1998;  98JP-00095448.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
PA   (KISH/) KISHIMOTO T.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K;
XX
DR   WPI; 1999-591042/50.
DR   P-PSDD; AAY39993.
XX
PT   CXCR4-potentiating agents and methods useful for inhibiting
PT   neovascularization, and treating solid cancers.
XX
PS   Disclosure; Page 48-49; 63pp; Japanese.
XX
CC   This sequence encodes the human CXCR4 protein. The invention relates to
```

```
CC   remedies inhibiting neovascularisation, remedies for solid cancer,
CC   remedies for diseases pathologically caused by neovascularisation and
CC   tissue repairing agents containing as the active ingredient a substance
CC   capable of potentiating CXCR4. Based on a finding that vascularisation is
CC   inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC   remedies inhibiting vascularisation which contain as the active
CC   ingredient a substance capable of potentiating CXCR4, remedies for solid
CC   cancer, remedies for diseases pathologically caused by neovascularisation
CC   and tissue repairing agents containing as the active ingredient a
CC   substance capable of potentiating CXCR4. It is also possible to establish
CC   methods for treatment with the use of these remedies
XX
SQ   Sequence 1588 BP; 396 A; 375 C; 334 G; 483 T; 0 U; 0 Other;

Query Match             75.0%;  Score 517.2;  DB 2;  Length 1588;
Best Local Similarity   85.8%;  Pred. No. 1e-136;
Matches  592;  Conservative    0;  Mismatches   83;  Indels   15;  Gaps    1;

Qy          1 CTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCTTCTGGGCAGTTGAT   60
              ||||||||||||||||||||||||||||| ||||| |||||| ||||||||||||||
Db        232 CTGCACCTGTCAGTGGCCGACCTCCCTCTTTGTCATCACGGTTCCCTTCTGGGCAGTTGAT  291

Qy         61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
              |||||||||||||| |||||||||||| ||| ||| || ||||||||||||||||||||
Db        292 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  351

Qy        121 GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
              |||||||||||||||||| ||||||||||||||||||||||| |||||||||||||||||
Db        352 GTCAACCTCTACAGCAGTGTCCTCATCCTGCCCTTCATCAGTCTGGACCGCTACCTGGCC  411

Qy        181 ATTGTCCACGCACCACCAGTCAAAGGCCAAGGAAACTGTGGCTGAAAAGGCAGTCTAT  240
              |||||||||||||  ||  ||||| ||||||||||||||||||| |||||| ||||||||
Db        412 ATCGTCCACGCACCAACAGTCAGAGCGCCAAGGAAGCTGTTGGCTGAAAAAGGTGGTCTAT  471

Qy        241 GTGGGCGTGTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
              ||||||||||||||||||||||||| |||||| ||||| |||||||||||||||| ||
Db        472 GTTGGGGTGTGATCCTGCCCTCCCGCTGCTGACTATCTTCATCTTTGCCAACG--  529

Qy        301 AGCCAGGGGGACATGCAGGGGATGACAGGTACATCTGTGACCGGCTTTACCCCGAT  360
              |||||||||||| ||||||||||||||||| |||||||||||||||||||||||||
Db        530 --------TCAGTGAGGCAGATGACAGAGATATCTGTGACCGGCTTCTACCCCAAT  576

Qy        361 AGCCTGTGGATGGTGTTTCAATTCAGCAGATATAAATGGTGGGTCTCATCCTGCCCGGC  420
              |||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||
Db        577 GACTTGTGGGGTGGTTGTGTTCCAGTTCAGCACACATGTTGGCCTTATCCTGCCTGGT  636

Qy        421 ATCGTCATCCTCTGTACTGACTCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC  480
              ||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||
Db        637 ATTGTCATCCTGTCCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC  696

Qy        481 CAGAAGCGCAGGCCCTCAAGACGCAGCAGTCATCCTCATCCTCCTAGCTTCTTGCCTGCTGG  540
              |||||||||||||||||||||||||| ||||||||| |||||||||||||||||||||||
Db        697 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTCCTGGCTTTCTTCGCCTGTTGG  756

Qy        541 CTGCCATATTATGTGGGATCAGCATCAGCATCCTTTTTGGGAGTCATCAAGCAA  600
              ||||| |  | ||| |||||||||||||| |||||||| |||||||||||||||
Db        757 CTGCCTTACTACATTGGGATCAGCATTGGATCCTTCTCCTGGAAATCATCAAGCAA  816

Qy        601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC  660
              ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        817 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC  876

Qy        661 TTCCACTGTGCCTGAACCCCATCCTCTAT  690
              ||||||||||||||||||||||||||||||
Db        877 TTCCACTGTGTGTCTGAACCCCATCCTCTAT  906
```

Search completed: May 17, 2004, 06:37:05
Job time : 298.894 secs

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:52:09 ; Search time 55.2446 Seconds
                              (without alignments)
                       6931.281 Million cell updates/sec

Title: US-09-367-052-3
Perfect score: 690
Sequence: 1 ctgcacctgtcagtggctga.........gcctgaacccatcctctat 690

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : Issued Patents NA:*
            1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	517.2	75.0	1225	4	US-09-016-434-1235	Sequence 1235, Ap
2	517.2	75.0	1317	4	US-08-153-848-45	Sequence 45, Appl
3	517.2	75.0	1317	3	US-09-299-843A-45	Sequence 45, Appl
4	517.2	75.0	1317	4	US-09-088-337B-45	Sequence 45, Appl
5	517.2	75.0	1317	5	PCT-US93-11153-45	Sequence 45, Appl
6	517.2	75.0	1664	4	US-09-582-224A-5	Sequence 5, Appli
7	517.2	75.0	1664	4	US-09-023-655-1213	Sequence 1213, Ap
8	517.2	75.0	1679	4	US-09-517-605-14	Sequence 14, Appl
9	517.2	75.0	1737	1	US-08-202-056-4	Sequence 4, Appli
10	517.2	75.0	1737	1	US-08-076-093A-3	Sequence 3, Appli
11	517.2	75.0	1737	1	US-08-701-265-3	Sequence 3, Appli
12	517.2	75.0	1737	2	US-08-284-586-3	Sequence 3, Appli
13	517.2	75.0	1737	2	US-08-805-478-3	Sequence 3, Appli
14	517.2	75.0	1737	2	US-08-802-627A-3	Sequence 3, Appli
15	517.2	75.0	1737	2	US-08-801-238-3	Sequence 3, Appli
16	517.2	75.0	1737	2	US-08-801-228-3	Sequence 3, Appli
17	517.2	75.0	1737	3	US-09-104-296-3	Sequence 3, Appli
18	517.2	75.0	1737	5	PCT-US94-06380-2	Sequence 2, Appli
19	150	21.7	1119	4	US-09-170-496D-65	Sequence 65, Appl
20	150	21.7	1679	1	US-08-202-056-6	Sequence 6, Appli
21	150	21.7	1679	1	US-08-076-093A-5	Sequence 5, Appli
22	150	21.7	1679	1	US-08-701-265-5	Sequence 5, Appli
23	150	21.7	1679	2	US-08-284-586-5	Sequence 5, Appli
24	150	21.7	1679	2	US-08-805-478-5	Sequence 5, Appli
25	150	21.7	1679	2	US-08-802-627A-5	Sequence 5, Appli
26	150	21.7	1679	2	US-08-801-238-5	Sequence 5, Appli
27	150	21.7	1679	2	US-08-801-228-5	Sequence 5, Appli
28	150	21.7	1679	3	US-09-104-296-5	Sequence 5, Appli
29	150	21.7	1679	5	PCT-US94-06380-3	Sequence 3, Appli
30	150	21.7	2818	3	US-08-982-493-7	Sequence 7, Appli
31	150	21.7	2818	3	US-08-628-655-1	Sequence 1, Appli
32	148.4		1119	4	US-09-170-496D-199	Sequence 199, App
33	142.2		1176	5	PCT-US95-0303Z-2	Sequence 2, Appli
34	142.2		1883	1	US-08-202-056-2	Sequence 1, Appli
35	142.2		1933	1	US-08-076-093A-1	Sequence 1, Appli
36	142.2		1933	1	US-08-410-451-1	Sequence 1, Appli
37	142.2		1933	1	US-08-410-455-1	Sequence 1, Appli
38	142.2		1933	1	US-08-418-919-1	Sequence 1, Appli
39	142.2		1933	2	US-08-410-453A-2	Sequence 2, Appli
40	142.2		1933	1	US-08-701-265-1	Sequence 2, Appli
41	142.2		1933	2	US-08-410-454A-2	Sequence 2, Appli
42	142.2		1933	1	US-08-284-586-1	Sequence 1, Appli
43	142.2		1933	2	US-08-410-456A-2	Sequence 2, Appli
44	142.2		1933	2	US-08-805-478-1	Sequence 1, Appli
45	142.2		1933	2	US-08-802-627A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-016-434-1235
; Sequence 1235, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/016,434
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1235:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1225 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g189313
US-09-016-434-1235

Query Match   75.0%;   Score 517.2; DB 4; Length 1225;
Best Local Similarity 85.8%;   Pred. No. 5.7e-140;

Left column alignment

Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
         |||||||||||||||| ||||||| |||||||||| || |||||||||||||||||||
Db  257 CTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  316

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTCCATATCATCTACACT         120
         ||| ||||  ||||||||||||||||  ||                |||||||||
Db  317 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  376

Qy  121 GTCAACCTCTACAGCAGCGTTCCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCC   180
         |||||||||||||||| ||  |||||| ||||||  |||| ||| |||||||||| |||
Db  377 GTCAACCTCTACAGCAGTGTCCTTCATCCTTCAGTCTGGCCGCGCCACCTGGCC        436

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGGAAAAGGCAGTCTAT  240
         ||||||||||||||||||||||  |||||||||||| ||||||||||||||||||||||
Db  437 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGCTGGCTGAAAAGGTGGTCTAT  496

Qy  241 GTGGGGTCTGGATCCCAGCCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC     300
         |||||||||||||  |||| ||||||  ||||||||| ||||||||||||||||||
Db  497 GTTGGCGTCTGGATCCCTGCCCCTCCTGACTATTCCCGACTATCTTTGCCAACG--      554

Qy  301 AGCCAGGGGACATCAGTCAGGGGATGACAGTACATCTGTGACCGCCTTTACCCCGAT     360
         ||||                |||||||||||||||||||||||||| |||||||||
Db  555 ------------TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT   601

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGCGGTCTCATCCTGCCCGGC     420
         ||||| ||||||||||||||||||||||| |||||||  ||||||||||||||||||
Db  602 GACTTGTGGGTGGTTGTGTTCCAGTTCAGCACATCATTGCGCCTTATCCTGCCTGGT    661

Qy  421 ATCGTCATCCTCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC  480
         ||||||||||||||||| ||||| || ||||| |||||||||||| |||||||||||||
Db  662 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGGCAC  721

Qy  481 CAGAAGCGCCAAGGCCCTCAAGACGACAGTCATCATCCTTCTTGCCTCTGG         540
         |||||| |||||||||||||||||||||||| | |||||| ||||| |||||
Db  722 CAGAAGCGCCAAGGCCCTCAAGACCACAGTCATCATGCATCCTGGCTTGCCTCGTTGG    781

Qy  541 CTGCCATATATGTGGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAA         600
         |||||  |||||||||| || ||||| |||||||||||| ||||||||||||
Db  782 CTGCCTTACTACATTGGGGATCAGCATCGCACTCCTGGGAAATCATCAAGCAA    841

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTTCCATCACAGAGGCCCTCGCCTTC 660
         ||| ||||||||||||||||||||||| ||| ||||||||||||||||||||||||||||
Db  842 GGGTGTGAGTTTGAGAGCATTGGGCACACTGTGCACAAGTGGATTTCCATCAGCAGAGGGCCCTAGCTTTC  901

Qy  661 TTCCACTGTTGCCTGAACCCCATCCTCTAT            690
         |||||| ||||||||||||||||| ||||||
Db  902 TTCCACTGTTGCTGAACCCCATCCTCTAT           931
```

Right column alignment

Query Match 75.0%; Score 517.2; DB 1; Length 1317;
Best Local Similarity 85.8%; Pred. No. 6e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
         |||||||||||||||| ||||||| |||||||||| ||       ||||||||||||||
Db  387 CTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCTTCACGCTTCCCTTCTGGGCAGTTGAT  446

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCATATCATCTACACT    120
         ||| ||||  ||||||||||||||||  ||          |||||| |||||||||| |
Db  447 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTTCATGTCATCATCTACACA  506

Qy  121 GTCAACCTCTACAGCAGCGTTCCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCC    180
         |||||||||||||||| ||  |||||| ||||||  |||| ||| |||||| |||| |||
Db  507 GTCAACCTCTACAGCAGTGTCCTTCATCCTTCAGTCTGGCCGCGCCACCTGGCC        566

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGGAAAAGGCAGTCTAT  240
         ||||||||||||||||||||||  |||||||||||| ||||||||||||||||||||||
Db  567 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGCTGGCTGAAAAGGTGGTCTAT  626

Qy  241 GTGGGGTCTGGATCCCAGCCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC    300
         |||||||||||||  |||| ||||||  ||||||||| ||||| ||||||||||||||
Db  627 GTTGGGGTCTGGATCCCTGCCCCTCCTGACTATTCCCGACTTCATCTTTGCCAACG--    684

Qy  301 AGCCAGGGGGACATCAGTCAGGGGATGACAGTACATCAGGTACATCTGTGACCGCCCTTACCCCGAT 360
         ||||                |||||||||||||||||||   ||||||||||||||| ||||||||||
Db  685 ------TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT    731

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGCGGTCTCATCCTGCCCGGC    420
         ||||| ||||||||||||||  |||||||||| ||   ||  |||||||||||| ||||
Db  732 GACTTGTGGGTGGTTGTGTTCCAGTTCAGCACATCATTGCGCCTTATCCTGCCTGGT  791

Qy  421 ATCGTCATCCTCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC  480
         |||||||||||||||||| ||||| || ||||| |||||||||||| |||||||||||||
Db  792 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGGCCAC  851

Qy  481 CAGAAGCGCCAAGGCCCTCAAGACGACAGTCATCATCCTTCTTGCCTCTTTTGGGAGTCATCAAGCAA 540
         |||||| ||||||||||||||||||| ||||| | |||| |||||||||||||||||| || ||||||
Db  852 CAGAAGCGCCAAGGCCCTCAAGACCACAGTCATCATCGCCGCTTCTTCTTCGCCTGTTGG  911

Qy  541 CTGCCATATATGTGGGGATCAGCATCATCCTCTTGGGAGTGGTCATCAAGCAA        600
         |||||  |||||||||| || |||||||| |||| |||||| ||||||||||||
Db  912 CTGCCTTACTACATTGGGGATCATGGCATCAGCATCAGCCTGGAAAATCATCAAGCAA  971

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCTTC  660
```

```
Db    447 GCCGTGGCAAACTGGTACTTTGGAGAACTTCCTATGCAGCAGTTCCATGCATGTCATCATCTACACA  506
Qy    121 GTCAACCTTCACAGCAGCGTTCATCCTCGGCCTTCATCCTCGGCCTTCATCAGCCTGACCTGCC  180
Db    507 GTCAACCTTCACAGCAGTGTCCTTCATCCTCGGCCTTCATCCTCGGCACCTGGACCGCTACCTGCC -  566
Qy    181 ATTGTCCACGCCACCAACAGTCAAGGCCAAGAGAAACTGCTGCTGAAAAAGGCAGTCTAT  240
Db    567 ATCGTCCACGCCACCAACAGTCAAGAGAGGCCAAGAAGAAGTGTTGGCTGAAAAAGGTGGTCTAT  626
Qy    241 GTGGGCGTCTGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGAGCTC  300
Db    627 GTTGGCGTCTGGATCCCCTGCCCTCCTGCCTGTGACTATTCCGACTTCATCTTTGCCAACG -  684
Qy    301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGTACATCGTGACCGCTTACCCCGAT  360
Db    685 ----------TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCACCCCCAAT  731
Qy    361 AGCCTGTGGATGGTGGTGTTCAATTCAGCATACATAAATGGTGGGTCTCATCCTGCCCGC  420
Db    732 GACTTGTGGGTGGTGGTTGTGTTCCAGTTTCAGCACCATCATGTGGCCTATCCTGCCTTGT  791
Qy    421 ATGTCATCCTCTCCTGTTACTGCATCATCATCTCAAGCTGTCACACTCCAAGGGGCCAC  480
Db    792 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGGCCAC  851
Qy    481 CAGAGCGGCAAGGCCCTCAAGAGACGACAGTCATCCTCCATCCTCCTAGCTTCTTTGCCTGCTGTGG  540
Db    852 CAGAAGCGCAAGGCCCTCAAGACCACACCCATCATCCTCATCCTCTGCCTTCTTCGCCTTGTTGG  911
Qy    541 CTGCCATATTATGTGGGGATCAGCATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA  600
Db    912 CTGCCTTACTACATTGGGGATCAGCATCAGCATCCACTCCTCCTCCTGGAAATCATCAAGCAA  971
Qy    601 GGATGTGACTTCGAGAGCAGCATTGTGCACAAGAGTGGATCCCATCACAGAGGCCCTCGCCTTC  660
Db    972 GGGTGTGAGTTGGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC  1031
Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
Db   1032 TTCCACTGTTGTCTGAACCCCATCCTCTAT  1061
```

```
RESULT 4
US-09-088-337B-45
; Sequence 45, Application US/09088337B
; Patent No. 6345574
GENERAL INFORMATION:
 APPLICANT: Godiska, Ronald
 APPLICANT: Gray, Patrick W.
 APPLICANT: Schweikart, Vicki L.
 TITLE OF INVENTION: No. 6345574e1 Seven Transmembrane Receptors
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
  ADDRESSEE: Borun
  STREET: 6300 Sears Tower, 233 South Wacker Drive
  CITY: Chicago
  STATE: Illinois
  COUNTRY: USA
  ZIP: 60606
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/088,337B
  FILING DATE: 01-Jun-1998
  CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/153,848
```

```
Db    972 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGGATTTCCATCACCGCGGCCCTAGCTTTC  1031
Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
Db   1032 TTCCACTGTTGTCTGAACCCCATCCTCTAT  1061

RESULT 3
US-09-299-843A-45
; Sequence 45, Application US/09299843A
; Patent No. 6107475
GENERAL INFORMATION:
 APPLICANT: Godiska, Ronald
 APPLICANT: Gray, Patrick W.
 APPLICANT: Schweikart, Vicki L.
 TITLE OF INVENTION: No. 6107475e1 Seven Transmembrane Receptors
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
  ADDRESSEE: Borun
  STREET: 6300 Sears Tower, 233 South Wacker Drive
  CITY: Chicago
  STATE: Illinois
  COUNTRY: USA
  ZIP: 60606
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/299,843A
  FILING DATE:
  CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 09/088,337
  FILING DATE: 01-JUN-1998
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/153,848
  FILING DATE: 17-NOV-1993
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/977,452
  FILING DATE: 17-NOV-1992
 ATTORNEY/AGENT INFORMATION:
  NAME: Jill E. Uhl
  REGISTRATION NUMBER: 43,213
  REFERENCE/DOCKET NUMBER: 27866/32059B
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: (312) 474-6300
  TELEFAX: (312) 474-0448
  TELEX:
INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
  LENGTH: 1317 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
  NAME/KEY: CDS
  LOCATION: 201..1211
US-09-299-843A-45

Query Match     75.0%;  Score 517.2;  DB 3;  Length 1317;
Best Local Similarity   85.8%;  Pred. No. 6e-140;
Matches  592;  Conservative  0;  Mismatches  83;  Indels  15;  Gaps  1;

Qy    1 CTGCACCTGTCAGTGGCTGACCTCCCTTCTCATCACACTCCCCTTCTGGGCAGTTGAT  60
Db  387 CTGCACCTGTCAGTGGCGGCCGACCTCCCTTCTCCTTCTCATCACGCCTTCCCTTCTGGGGCAGTTGAT  446
Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT  120
```

FILING DATE: 17-NOV-1993
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6348574and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1317 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 201..1211
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-088-337B-45

Query Match   75.0%;   Score 517.2; DB 4; Length 1317;
Best Local Similarity 85.8%;   Pred. No. 6e-140;
Matches 592; Conservative   0; Mismatches   83; Indels   15; Gaps   1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT   60
           |||||||||||||||||||||| |||||||| |||||||||||||| || |||||||||
Db   387 CTGCACCTGTCAGTGGCCGACCTCCTCTTTTGTCATCACGCTTCCTTCTCGGGCAGTTGAT   446

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
           ||| ||||||||||||||||||||| |||||||| |||||||||||||||| |||||
Db   447 GCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGCAGTCCATGTCATCATCACACA   506

Qy   121 GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCC   180
           |||||||||||||| ||||||||||||||||| ||||||||||||||||| |||||
Db   507 GTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGCTACCTGGCC   566

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
           || ||||||||||||| |||||| ||||||||||| ||||| ||||||||| |||||
Db   567 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGGCTGAAAAGGTGGTCTAT   626

Qy   241 GTGGGCGTCTGGATCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC   300
           |||||||||||||||||| |||||||| |||||||||||||| |||||||||||||||
Db   627 GTTGGCGTCTGGATCCTGCCCCCTGCTGACTATTCCCGACTTCATCTTTGCCAACG--   684

Qy   301 AGCCAGGGGGGACATCAGTCAGGGGGATGACAGTACATCGTGACCGCCTTTACCCCGAT   360
           ||| | |||||| |||||||||||||||||||||||||||||||||||||
Db   685 --------TCAGTGAGGCAGATGACAGATATATCGTGACCGCTTCTCACCCCAAT   731

Qy   361 AGCCTGTGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC   420
           |||||| |||||| |||||||| |||||||||||||||||||||||||||||||
Db   732 GACTTGTGGGTGGTGTGTTCCAGTTTCAGCACCAATCATGGTTGGCCTTATCCTGCCCTGGT   791

Qy   421 ATCGTCATCCTCCTCCTGTTACTGCATCATCATCCTCAAGCTGTCACACTCCAAGGGCCAC   480
           ||| ||||| |||||| ||| ||||||||| |||||||| ||||||||||||
Db   792 ATTGTCATCCTGTCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC   851

Qy   481 CAGAAGGCAGGCCCTCAAGACGACGAGTCATCCTCTCATCCTCATCCTGGCTTCTTCTTGCCTTGCTGG   540
           ||||||||||||||| |||||||||| ||||||||||| ||||||||| ||||||||
Db   852 CAGAAGCGGCAAGGCCCTCAAGACCACAGACCTCATCCTCATCCTGGCTTTCTTGCCTGTTGG   911

Qy   541 CTGCCATATTATGTGGGGATCAGCATGCATCCGACTCCCTTTTGGGAGTCATCAAGCAA   600
           ||||||| ||||| ||||||||||||||||||||| ||||||||||||||||||
Db   912 CTGCCTTACTACATTGGGATCAGCATGCATCCGACTCCCTGGAAATCATCAAGCAA   971

Qy   601 GGATGTGACTTCGAGAGACATTGTGCACAAGTGGATCTCCATCACAGAGAGCCCTCGCCTTC   660
           ||||||||||||||||||| ||||| |||||||||||||||||||||| ||||||||
Db   972 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGGATTTCCATCACCGGAGCCCTAGCTTTC   1031

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT   690

Db   1032 TTCCACTGTGTGCTGAACCCCATCCTCTAT   1061
           |||||||||| ||||||||||||||||||||

RESULT 5
PCT-US93-11153-45
; Sequence 45, Application PC/TUS9311153
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: Novel Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11153
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 201..1211
PCT-US93-11153-45

Query Match   75.0%;   Score 517.2; DB 5; Length 1317;
Best Local Similarity 85.8%;   Pred. No. 6e-140;
Matches 592; Conservative   0; Mismatches   83; Indels   15; Gaps   1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT   60
           |||||||||||||||||||||| |||||||| |||||||||||||| || |||||||||
Db   387 CTGCACCTGTCAGTGGCCGACCTCCTCTTTTGTCATCACGCTTCCTTCTCGGGCAGTTGAT   446

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
           ||| ||||||||||||||||||||| |||||||| |||||||||||||||| |||||
Db   447 GCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGCAGTCCATGTCATCATCACACA   506

Qy   121 GTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCC   180
           |||||||||||||| ||||||||||||||||| ||||||||||||||||| |||||
Db   507 GTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGCTACCTGGCC   566

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
           || ||||||||||||| |||||| ||||||||||| ||||| ||||||||| |||||
Db   567 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGGCTGAAAAGGTGGTCTAT   626

Qy   181 ATTGTCCACGCCACCAACAGTGCACAAGTGGATCTCCATCACAGAGGAAACTGCTGGCTGCGCTCTAT   240
           ||| || |||||||| ||||| |||||||||| ||| |||||||
Db   567 ATCGTCCACGCCACCAACAGTGCACAAGTGCACAAGTGGGATTTCCATCACCGGAGCCGCTAGCTTTC   626

Qy   567 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGGCTGAAAAGGTGGTCTAT   690

Qy 241 GTGGGGGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
        ||||||||||||| |||||||||||||||||| ||||||||||||| |||||||
Db 627 GTTGGGGCTGTCGGATCCCGCCCCTCCTGCTGACTATTCCCGACTTCATCTTTGCCAACG-- 684

Qy 301 AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 360
                 ||||||||||||||||||||||||||||||||
Db 685 --------TCAGTGAGGCAGATGACAGATATATCTGTGACCGCGTTCTCACCCCCAAT 731

Qy 361 AGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCGGC 420
        ||||| ||| |||||| ||||||||||||
Db 732 GACTTGTGGGTGGGTGTTCCAGTTTGGCCTTATCCTGCCTGGT 791

Qy 421 ATCGTCATCCTCTCGTTACTGATCATCATCTCCTAAGCTGTCACACTCCAAGGGCCAC 480
        |||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 792 ATTGTCATCCTGCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 851

Qy 481 CAGAAGCGCAAGGCCCTCAAGACGACCAGTGTCTTCCTCATCCTCAGCTTCTTGCCTGCTGG 540
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 852 CAGAAGCGCAAGGCCCTCAAGACCACCAGTGTCATCCTCGGCTTTCTTCGCCTGCTGTTGG 911

Qy 541 CTGCCATTATGTGGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAA 600
        |||||||||||||||||||||||||||||||||||||||||
Db 912 CTGCCTTACTACATTGGGATCAGCATGACGACCACAGTGATCATCATCAAGCAA 971

Qy 601 GGATGTGACTTCGAGACGCATTGTGCACAAGTGGATCCTCCATCACAGAGGCCTCGCCTTC 660
        |||||||||||||||||||||||||||||||||||||||
Db 972 GGGTGTGAGTTTGAGACACTGTGCACAAGTGGATTTCCATCACCGGAAATCATCAAGCAA 1031

Qy 661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
        ||||||||||||||||||||||||||||||
Db 1032 TTCCACTGTTGTCTGAACCCCATCCTCTAT 1061

RESULT 6
US-09-582-224A-5
; Sequence 5, Application US/09582224A
; Patent No. 6429308
; GENERAL INFORMATION:
;   APPLICANT: IIJIMA, Osamu
;   APPLICANT: GOTO, Takeshi
;   APPLICANT: SHIMADA, Takashi
;   TITLE OF INVENTION: HIV Infection Inhibitors
;   FILE REFERENCE:
;   CURRENT APPLICATION NUMBER: US/09/582,224A
;   CURRENT FILING DATE: 2000-07-21
;   PRIOR APPLICATION NUMBER: PCT/JP99/06534
;   PRIOR FILING DATE: 1999/11/24
;   NUMBER OF SEQ ID NOS: 5
;   SOFTWARE: Microsoft Word
; SEQ ID NO 5
;   LENGTH: 1664
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: cDNA of CXCR4
US-09-582-224A-5

Query Match   75.0%; Score 517.2; DB 4; Length 1664;
Best Local Similarity   85.8%; Pred. No. 6.7e-140;
Matches   592; Conservative   0; Mismatches 83; Indels 15; Gaps   1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 308 CTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCACGCTCCCTTCTGGGCAGTTGAT 367

Qy   61 GCCATGGCTGACTGGTACTTTGGGGAAATTTTGTGTAAGGCTGTCCATATCATCTCACACT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 368 GCCGTGGCAAACTGGTACTTTGGGACTTCCTATGCAAGGCAGTCCATGTCATCATCCACACA 427

Qy 121 GTCAACCTCTACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACACGGTACCTCGCC 180
        ||||||||||||||||||||||||||||||||||||||||||||
Db 428 GTCAACCTCTACAGCAGCAGTGTCCTCATCAGTCTGGCCTTCATCAGTCTGGACACCGCTACCTGGCC 487

RESULT 7
US-09-023-655-1213
; Sequence 1213, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
;   APPLICANT: Cocks, Benjamin G.
;   APPLICANT: Susan G. Stuart
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;   TITLE OF INVENTION: EXPRESSION
;   NUMBER OF SEQUENCES: 1508
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/023,655
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
;   TELECOMMUNICATION INFORMATION:

Qy 181 ATTGTCCACGCACCAACAGTCAAAGGCCAAGAAAACTGTGCTGCTGAAAAGGCAGTCTAT 240
        |||||||||||||||||||||||||||||||||||||||||
Db 488 ATCGTCCACGCCAGCCAACACAGTCAGAGGCCAAGGAGAGTCGTGTTGGCTGAAAAGGTCGTCTAT 547

Qy 241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
        |||||||||||||||||||||||||||||||||||||
Db 548 GTTGGCGTCTGGATCCCGCCCCTCCTGCTGACTATTCCCGACTTCATCTTTGCCAACG- 605

Qy 301 AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 360
                 ||||||||||||||||||||||||||||||||
Db 606 --------TCAGTGAGGCAGATGACAGATATATCTGTGACCGCGTTCTCACCCCCAAT 652

Qy 361 AGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCGGC 420
        ||||| ||| |||||| ||||||||||||
Db 653 GACTTGTGGGTGGGTGTTCCAGTTTGCAGCACATCATGGTTGGCCTTATCCTGCCTGGT 712

Qy 421 ATCGTCATCCTCTCGTTACTGCATCATCATCTCCTAAGCTGTCACACTCCAAGGGCCAC 480
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 713 ATTGTCATCCTGCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 772

Qy 481 CAGAAGCGCAAGGCCCTCAAGACGACAGTGCTCATCCTCATCCTCAGCTTCTTGCCTGCTGG 540
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 773 CAGAAGCGCAAGGCCCTCAAGACCACCAGTGTCATCCTCGGCTTTCTTCGCCTGCTGTTGG 832

Qy 541 CTGCCATTATGTGGGGATCAGCATCAGCACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
        |||||||||||||||||||||||||||||||||||||||||
Db 833 CTGCCTTACTACATTGGGATCAGCATGACGATCAGCATGACTCCTTCATCCTTCATCCTCCTGGAAATCATCAAGCAA 892

Qy 601 GGATGTGACTTCGAGACGCATTGTGCACAAGTGGATCCTCCATCACAGAGGCCCTCGCCTTC 660
        |||||||||||||||||||||||||||||||||||||||
Db 893 GGGTGTGAGTTTGAGACACTGTGCACAACACTGTGCACAAGTGGATTTCCATCACCGGAGGCCCCTAGCTTTC 952

Qy 661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
        ||||||||||||||||||||||||||||||
Db 953 TTCCACTGTGTCTGAACCCCATCCTCTCTAT 982

Query Match   75.0%;   Score 517.2;   DB 4;   Length 1679;
Best Local Similarity 85.8%;   Pred. No. 6.7e-140;
Matches 592; Conservative   0; Mismatches 83; Indels 15; Gaps   1;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT 60
        |||||||||||||||| ||||||||| ||||||||||||||| |||||||||||||||||||
Db  320 CTGCACCTGTCAGTGGCCGACCTCCTCTCTTTGTCATCACGCTTCCCTTCTCGGGCAGTTGAT 379

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
        ||| |||| || ||||||||||||||| ||   |  ||||||| ||||| |||||||||
Db  380 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA 439

Qy  121 GTCAACCTCTACAGCAGCGTTCCTCATCCTCGGCCTTCATCGGCCTTCATCCTCGGCCTACCTCGCC 180
        |||||||||||||||||| ||||||||||||||||||||||| |||||||||||||||||||||
Db  440 GTCAACCTCTACAGCAGTGTCCTCATCCTCGGCCTTCATCCTCAGTGTCCTTCATCAGTGTCCTACCTCGCC 499

Qy  181 ATTGTCCACGCCACCAACGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
        || |||||||||||||||| ||| |||||||||| ||||| ||||||||| |||||||
Db  500 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGTGGTCTAT 559

Qy  241 GTGGGCGGTCTGGATCCGAGCCCTCCTCCTGACTATACCTCATCTTTGCGACGTC 300
        || ||||| |||||||||  |||||||||||||||||| |||| ||||||| ||
Db  560 GTTGGCGCTCTGGATCCGCTCCTCCTGACTATTCCGACTTCATCTTTGCCAACG-- 617

Qy  301 AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCGCCTTTACCCCGAT 360
                        |||||| ||||| ||||||||||| ||||||||||| ||| |||||
Db  618 ---------TCAGTGAGGCAGATGACAGATATATCGTGACCGGCTTCTACCCCAAT 664

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTCGCCCGGC 420
        | ||||||| ||||| ||||||| |||||||| ||| ||||| ||| ||||||| ||||
Db  665 GACTTGTGGGTGGTTGTGTTCCAGTTTCAGCACATCATGGTTGGCCTTATCCTGCCTGGT 724

Qy  421 ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGGCCAC 480
        || |||||||| ||| | ||| ||| || |||||  ||||||||||||||||||||||||
Db  725 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGGCCAC 784

Qy  481 CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTCTTTTGGCTGCTGG 540
        |||||||||||||||||||||| ||| ||||||||||||||||||| ||||| ||
Db  785 CAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTCTGGCTTTCTTCGGCCTGTTGG 844

Qy  541 CTGCCATATTATGTGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
        ||||||| |  ||||| |||||||| ||| ||||||||||| |||||| ||||| ||||
Db  845 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA 904

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
        |||||||||||||| || |||||||| ||||||||  ||||||||||| ||||||||||
Db  905 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGCACAAGTGGCACCACCCGGCCCCTAGCTTTC 964

Qy  661 TTCCACTGTTGCCTGAACCCATCCTCTAT 690
        |||||||||||||||||||||||||||||
Db  965 TTCCACTGTTGTCTGAACCCCATCCTCTAT 994
```

Query Match   75.0%;   Score 517.2;   DB 4;   Length 1664;
Best Local Similarity 85.8%;   Pred. No. 6.7e-140;
Matches 592; Conservative   0; Mismatches 83; Indels 15; Gaps   1;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT 60
        |||||||||||||||| ||||||||| ||||||||||||||| |||||||||||||||||||
Db  308 CTGCACCTGTCAGTGGCCGACCTCCTCTCTTTGTCATCACGCTTCCCTTCTCGGGCAGTTGAT 367

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
        ||| |||| || ||||||||||||||| ||   |  ||||||| ||||| |||||||||
Db  368 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA 427

Qy  121 GTCAACCTCTACAGCAGCGTTCCTCATCCTCGGCCTTCATCGGCCTTCATCCTCGGCCTACCTCGCC 180
        |||||||||||||||||| ||||||||||||||||||||||| |||||||||||||||||||||
Db  428 GTCAACCTCTACAGCAGTGTCCTCATCCTCGGCCTTCATCCTCAGTGTCCTTCATCAGTGTCCTACCTCGCC 487

Qy  181 ATTGTCCACGCCACCAACGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
        || |||||||||||||||| ||| |||||||||| ||||| ||||||||| |||||||
Db  488 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGTGGTCTAT 547

Qy  241 GTGGGCGGTCTGGATCCGAGCCCTCCTCCTGACTATACCTCATCTTTGCGACGTC 300
        || ||||| |||||||||  |||||||||||||||||| |||| ||||||| ||
Db  548 GTTGGCGCTCTGGATCCGCTCCTCCTGACTATTCCGACTTCATCTTTGCCAACG-- 605

Qy  301 AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCGCCTTTACCCCGAT 360
                        |||||| ||||| ||||||||||| ||||||||||| ||| |||||
Db  606 ---------TCAGTGAGGCAGATGACAGATATATCGTGACCGGCTTCTACCCCAAT 652

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTCGCCCGGC 420
        | ||||||| ||||| ||||||| |||||||| ||| ||||| ||| ||||||| ||||
Db  653 GACTTGTGGGTGGTTGTGTTCCAGTTTCAGCACATCATGGTTGGCCTTATCCTGCCTGGT 712

Qy  421 ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGGCCAC 480
        || |||||||| ||| | ||| ||| || |||||  ||||||||||||||||||||||||
Db  713 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGGCCAC 772

Qy  481 CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTCTTTTGGCTGCTGG 540
        |||||||||||||||||||||| ||| ||||||||||||||||||| ||||| ||
Db  773 CAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTCTGGCTTTCTTCGGCCTGTTGG 832

Qy  541 CTGCCATATTATGTGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
        ||||||| |  ||||| |||||||| ||| ||||||||||| |||||| ||||| ||||
Db  833 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA 892

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
        |||||||||||||| || |||||||| ||||||||  ||||||||||| ||||||||||
Db  893 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCCTAGCTTTC 952

Qy  661 TTCCACTGTTGCCTGAACCCATCCTCTAT 690
        |||||||||||||||||||||||||||||
Db  953 TTCCACTGTTGTCTGAACCCCATCCTCTAT 982
```

; APPLICANT: Chuntharapai, Anan
; APPLICANT: Hebert, Caroline
; APPLICANT: Kim, Kyung Jin
; APPLICANT: Lee, James
; TITLE OF INVENTION: Antibodies to Human IL-8 Type B Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/202,056
;   FILING DATE: 25-FEB-1994
;   CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/677211
;   FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B.
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: 706P3
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1737 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-08-202-056-4

Query Match   75.0%;   Score 517.2; DB 1; Length 1737;
Best Local Similarity 85.8%;   Pred. No. 6.8e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

Qy   1 CTGCACCTGCAGTGGCTGACCTCCTCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
            |||||||||||||||||| |||||||||||||||| |||||||||||||||||||||||
Db   322 CTGCACCTGTTCAGTGGCGTCAGTGGCCGGACCTCCTCCTTTGTCATCACGGTTCCCTTCTGGGCAGTTGAT 381

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTGTGTAAGGCTGTCCATATCATCTACACT 120
            ||||||||||||||||||||||||||||||| |||||||||||||| |||||||||||
Db   382 GCCGTTGGCAAACTGGTACTTTTGGGAACTTCCTATGCAAGGCAGTCATCATCTACACA 441

Qy   121 GTCAACCTCTACAGACAGCGGTTCATCACGGTTCATCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
            ||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   442 GTCAACCTCTACAGACAGTGTCCTCAACAGCAGGCCTTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCC 501

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
            |||||||||||||||||||||||||||||||
Db   502 ATCGTCCACGCCACCAACAGTCAGAGGGCCAAGGGAAGCTGCTGTTGGCTGAAAAGGTGGTCTAT 561

Qy   241 GTGGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
            ||||||||||||||||||||
Db   562 GTTGGGGTCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCATCTTTGCCAACG-- 619

Qy   301 AGCCAGGGGGACATCAGTCAGGGGATGACAGAGTACATCTGTGACCCGCCTTACCCCGAT 360
            |||||||||||||||||||||||
Db   620 ----------TCAGTGAGGCAGATGACAGATATATCTGTGACCCGCCTTCACCCCCAAT 666

Qy   361 AGCCTGTGGATGAGTGGTGTTCAATTCCAGCATAATGTGGGTCTCATCCTGCCCCGGC 420
            | |||||||||| |||| ||||| |||||||||||||
Db   667 GACTTGTGGGTGGTGTGTTCCAGTTCCAGTTCCAGTTCAGCACACATCATGGTTGGCCTTATCCTTGCCTTATCCTTGCCCTGGT 726

RESULT 10
US-08-076-093A-3
; Sequence 3, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:
;   APPLICANT: Chuntharapai, Anan
;   APPLICANT: Lee, James
;   APPLICANT: Hebert, Caroline
;   APPLICANT: Jin Kim, K.
;   TITLE OF INVENTION: Antibodies to Human PF4A Receptors
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/076,093A
;   FILING DATE: 11-Jun-1993
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/810782
;   FILING DATE: 19-DEC-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/677211
;   FILING DATE: 29-MAR-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: 706P2
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
;   INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1737 nucleotides
;   TYPE: Nucleic Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Linear
US-08-076-093A-3

Query Match   75.0%;   Score 517.2; DB 1; Length 1737;
Best Local Similarity 85.8%;   Pred. No. 6.8e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

Qy   421 ATCGTCATCCTCCTCCTGTTACTGCATCATCATCTCAAGCTGTCACACTCCAAGGGCCAC 480
            |||||||||||||| |||||||||||||||||||||||||||
Db   727 ATTGTCATCCTGTCCTGCTATTGCATTGCATTGCATTCATCATCATCCAAGCTGTCACACTCCAAGGGCCAC 786

Qy   481 CAGAAGCGGCAAGGCCCTCAAGACAGACAGTCATCCTATCCTCAGCTTCTTTGCCTGCTGG 540
            ||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   787 CAGAAGCGGCAAGGCCCTCAAGACCACAGTCATCCTATCCTGGCTTTCTTCGCCTGTGTTGG 846

Qy   541 CTGCCATATTATGTGGGGATCAGCATGCAGGACTCCTTCATCCTTTGGGAGTCATCAAGCAA 600
            |||| || |||| ||||| ||||||||||||||||||||
Db   847 CTGCCTTACTACATTGGATCAGCATGCAGCATCGGAATCATCAAGCAA 906

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGATCCTCATCACAGAGCCCTCGCCTTC 660
            |||| |||| || ||| |||||||||||||||||||||||||||||||
Db   907 GGGTGTGAGTTTGAGAACACTGTGTCACAAGTGATTCATCACCGAGCCGAGCCCTAGCTTTC 966

Qy   661 TTCCACGTGTGCCTGAACCCCATCCTCCTCAT 690
            ||||||||||||||||||||||||||
Db   967 TTCCACTGTTGTCTGAACCCCATCCTCTAT 996

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/701,265
  FILING DATE: 22-AUG-1996
  CLASSIFICATION: 424
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/076093
  FILING DATE: 11-Jun-1993
  APPLICATION NUMBER: 07/810782
  FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/677211
  FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Love, Richard B
  REGISTRATION NUMBER: 34,659
  REFERENCE/DOCKET NUMBER: 706P2
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415/225-5530
  TELEFAX: 415/952-9881
  TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1737 nucleotides
    TYPE: Nucleic Acid
    STRANDEDNESS: Single
    TOPOLOGY: Linear
US-08-701-265-3:

Query Match   75.0%;   Score 517.2;   DB 1;   Length 1737;
Best Local Similarity 85.8%;   Pred. No. 6.8e-140;
Matches   592;   Conservative   0;   Mismatches   83;   Indels   15;   Gaps   1;

```
Qy      1 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
          |||||||||||||||  ||||||| | ||||||||||| || |||| ||||||||||||||
Db    322 CTGCACCTGTCAGTGCCGACCTCCTCCTTGTCATCACGCTTCCCTCTGGGCAGTTGAT 381

Qy     61 GCCATGCTGACTGGTACTTGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
          ||| ||| |||||| ||||| ||  ||||| || ||||| |||| |||| ||||||||
Db    382 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA 441

Qy    121 GTCAACCTCTACAGCACGTTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCC 180
          |||||||||||||||| ||| | ||| || |||||||||| |||| ||||| ||| ||
Db    442 GTCAACCTCTACAGCAGTGTCCTCATCCTCGGCCTTCATCAGTCTGGACCGGCTACCTCGGCC 501

Qy    181 ATTGTCCACGCCACCAACAGTCGGCTGCTGAAAAGGCAGTCTAT 240
          || |||||||||| ||||||||| | || |||||| ||||||||
Db    502 ATCGTCCACGCCGCCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAAGTGGTCTAT 561

Qy    241 GTGGGCGTCTGGATCCCAGCCCTCCTGCTGCTGCGACTATACCTGACTTCATCTTTGCCGACGTC 300
          || ||||||||||||||| ||||||||||||||| |||||||| ||||||||||| |||||| 
Db    562 GTTGGGCGTCTGGATCCCGCCCTCCTGCTGCTGCGACTATTCCCGACTTCATCTTTGCCAACG-- 619

Qy    301 AGCCAGGGGACATCAGTCAGGGGAGTACAGTCATCGTGTACCCGGAT 360
          |||||||  |||||| || |||| | || ||||| |||| |||||||
Db    620 ------------TCAGTGAGGCAGCAGTGCACAAGTGCACACTCCCAGCTTTCTTTGCCTGCTGG 666

Qy    361 AGCCTGTGATGGTGGTTGTGTTCAATTCAGCTTCAGCACATAATGGTGGCTTATCCTGCCGGC 420
          |||||||||||||||||||||||| |||  |||||||| |||||| ||||  ||||| | ||
Db    667 GACTTGTGTGGGTGTTGTTGCATTCAGCACATCATGTCACTATTACATCCCTGGCTTATCCTGCCGGT 726

Qy    421 ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 480
          || |||||||| || |||| |||| ||| || |||||||||||||||||| ||||||
Db    727 ATTGTCATCCTGTCCTGTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 786

Qy    481 CAGAAGCGCAAGGCCCTCAAGACGAGTACAGTCATCCTGTTGCTGG 540
          |||||||||||||||||| ||||| ||||| | ||||||| ||||
Db    787 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCATCCTGGCTTTGCTTGCTGG 846

Qy    541 CTGCCATATTATGTGGGATCACGTTCACTCCTTTTTGGGAGTCATCAAGCAA 600
          |||| |||| ||||||| || || ||| ||||| |||| ||||||||||||
Db    847 CTGCCTTACTACTACACATTGGGATCACGACTCCTTTCTCCTGAAATCATCAAGCAA 906

Qy    601 GGATGTGACTTGAGAGCATTGTGCACAAGTGGATCCTCATCACAGAGGCCCTCGCCTTC 660
          |||||||||||||||||| ||||||| ||| || || || ||||||||||| ||||||
Db    907 GGGTGTGAGTTTGGAGAACACTGTGCACAAGTGGATTCCATCACCACCGGAGCCCTAGCTTTC 966

Qy    661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
          ||||| |||||||||||||||||||||||
Db    967 TTCCACTGTTGCTGAACCCCATCCCTCTAT 996
```

RESULT 11
US-08-701-265-3
; Sequence 3, Application US/08701265
; Patent No. 5776457
GENERAL INFORMATION:
  APPLICANT: Chuntharapai, Anan
  APPLICANT: Lee, James
  APPLICANT: Hebert, Caroline
  APPLICANT: Jin Kim, K.
  TITLE OF INVENTION: Antibodies to Human PF4A Receptors
  NUMBER OF SEQUENCES: 6
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Genentech, Inc.
    STREET: 460 Point San Bruno Blvd
    CITY: South San Francisco
    STATE: California
    COUNTRY: USA
    ZIP: 94080
  COMPUTER READABLE FORM:
    MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: WinPatin (Genentech)

Db   442 GTCAACCTTCACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGCTACCTGGCC 501

Qy   181 ATTGTCCACGCACCACAGTCAAAGGCCAAGGAACTGCTGGCTGAAAAGGCAGTCTAT 240
         |||||||||||||||| |||||||||||||| |||||||||||||| ||||| |||
Db   502 ATCGTCCACGCACCACAACAGTCAGAGGCCAAGGAGCTGTTGGCTGAAAAGGTGGTCTAT 561

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTCGTGCTATACCTGACTCATCTTTGCCGACGTC 300
         ||||||||||||||||||||||||||||| ||||||||||| |||||||||||||| --
Db   562 GTTGGCGTCTGGATCCCGTGCTCCTCGTGCTACTATTCCCGACTATTCCATCTTTGCCAACG-- 619

Qy   301 AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCGCCTTACCCGAT 360
Db   620 ----------TCAGTGAGGCAGATGACAGATATATCTGTGACCGCCTTCTCACCCCAAT 666

Qy   361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAAATGGTGGGTCTCATCCTGCCCGGC 420
         |||||||||||||||||||||||||||||||| ||||||||| ||||||||||||||
Db   667 GACTTGTGGGTGGTGTTGTTCAGTTTCAGCATCATGTGGCCTTATCCTGCCTGGT 726

Qy   421 ATCGTCATCCTCCTGTTACTGATCATCATCTTCTAAGCTGTCACACTCCAAGGGCCAC 480
         |||||||||||||||||||| |||||||||||||| |||||||||||||||||||||||
Db   727 ATTGTCATCCTGCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 786

Qy   481 CAGAGGCGCAAGGCCCTCAAGACGCAGTCAGTCACCTCCTCATCACCTGGCTTTCTTTCGCTCTGTG 540
         |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   787 CAGAGGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTGGCTTTCTTCTTCGCCTGTTGG 846

Qy   541 CTGGCCATATTATGTGGGGATCAGCATCAGCATCGACTCCCTTCATCCTTTGGGAGTCATCAAGCAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847 CTGCCTTACTACATTGGGATCAGCATCAGCATCGACTCCCTCCTGGAAATCATCAAGCAA 906

Qy   601 GGATGGACTTCGAGAGCATTGTGCACAGTGGATCTCCATCACAGAGAGCCCTCGCCTTC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGTGGATTTCCATCACCGAGGCCCTAGCTTTC 966

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
         ||||||||||||||||||||||||||||||
Db   967 TTCCACTGTTGTCTGAACCCCCATCCTCTAT 996

RESULT 13
US-08-805-478-3
; Sequence 3, Application US/08805478
; Patent No. 5874543
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,478
; FILING DATE: 25-Feb-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:

Db   907 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGTGGATTTCCATCACCGAGGCCCTAGCTTTC 966
         || |||| |||||||||||||||||||||||||||||||||| ||||| || ||| |||
Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
Db   967 TTCCACTGTTGTCTGAACCCCCATCCTCTAT 996

RESULT 12
US-08-284-586-3
; Sequence 3, Application US/0284586
; Patent No. 5840856
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,586
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,093A
; FILING DATE: 11-Jun-1993
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 nucleotides
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-284-586-3

Query Match   75.0%; Score 517.2; DB 2; Length 1737;
Best Local Similarity   85.8%; Pred. No. 6.8e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps   1;

Qy   1 CTGCACCTGCTGCAGTGGCTGACCTCCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
         ||||||||||||||||||||||||| ||||||||||||||||||| ||||||||||||
Db   322 CTGCACCTGCTGCAGTGGCCGACCTCCTTCTGTCATCACACGGCTTCCCTTCTGGGCAGTTGAT 381

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
         ||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   382 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTCACACA 441

Qy   121 GTCAACCTTCACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGCTACCTGGCC 180
         ||||||||||||||||||||||| |||||||||||||||||| ||||||||

Left column

```
        APPLICATION NUMBER: 07/810782
        FILING DATE: 19-DEC-1991
     ATTORNEY/AGENT INFORMATION:
        NAME: Love, Richard B.
        REGISTRATION NUMBER: 34,659
        REFERENCE/DOCKET NUMBER: P0706P2P1C1
     TELECOMMUNICATION INFORMATION:
        TELEPHONE: 415/225-5530
        TELEFAX: 415/952-9881
        TELEX: 910/371-7168
  INFORMATION FOR SEQ ID NO: 3:
     SEQUENCE CHARACTERISTICS:
        LENGTH: 1737 base pairs
        TYPE: Nucleic Acid
        STRANDEDNESS: Single
        TOPOLOGY: Linear
US-08-805-478-3
```

```
Query Match          75.0%;   Score 517.2;   DB 2;   Length 1737;
Best Local Similarity 85.8%;   Pred. No. 6.8e-140;
Matches  592;   Conservative  0;   Mismatches  83;   Indels  15;   Gaps  1;

Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
          |||||||||||||||  ||||||||| | ||||||||||| ||| ||||||||||||||||
Db  322 CTGCACCTGTCAGTGGCCGACCTCCTTCCCTTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  381

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
          |||  ||| ||||||||||||||||||   |  ||  |||||||||| | ||||||||| |
Db  382 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  441

Qy  121 GTCAACCTCTACAGCAGCGGTTCCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
          ||||||||||||||||| || |||||||||||||||||| | |||||||| ||||| |||
Db  442 GTCAACCTCTACAGCAGTGTCCTTCATCCTGGCCTTCATCAGTCTGGACCGCTACCTGGCC  501

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
          || |||||||||||||||||||| | ||||||||||| | | ||||||||||| |||||||
Db  502 ATCGTCCACGCCACCAACAGTCAGGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT  561

Qy  241 GTGGGCGTCTGGATCCCCAGCCCTCCTCCTGACTATACCTGGTCATCTTTGCCGACGTC   300
          || |||||||||||||||  |||||||  | ||| || || ||||||||||||| |||||
Db  562 GTTGGCGTCTGGATCCCCTGCCCTCCTGCTATTCCCGACTTCATCTTTGCCAACG--    619

Qy  301 AGCCAGGGGACATCAGTCAGGGGATGACAGTGACATCTGTGACCGGCTTTACCCCGAT   360
                ||||| ||||| |||  ||| |||||  |  || ||||||||| || |||||| |
Db  620 -------TCAGTGAGGCAGATGACAGATATATCTGTGACCGGCTTCTACCCCAAT    666

Qy  361 AGCCTGTGATGGTGGTGTTTCAATTCCAGCATAATGAGGTGGGTCTCATCCTGCCGGC   420
          |  ||||| |||||||| ||||| ||||| |||| | ||||||| || ||||||||| |
Db  667 GACTTGTGGTGGTTTGTGTTCCAGTTTCACACATCATGGTGGGCTTTATCCTGCCGGT   726

Qy  421 ATCGTCATCCTGCTCCTGTTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC  480
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  727 ATTGTCATCCTGCTCCTGTTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC  786

Qy  481 CAGAAGCGCAAGGCCCTCAAGACGCACAGTACTTCTTGCTTGCTGG  540
          |||||||||||||||||||||||||||||||| || || ||||||
Db  787 CAGAAGCGCAAGGCCCTCAAGACCACCAGTACCACCAGTGCTTGG  846

Qy  541 CTGCCATATTATGTGGGGATCAGCACTCCCTTTGGGAGTCATCAAGCAA  600
          | ||| | | || ||||| ||||||||| || |||||| |||||||||
Db  847 CTGCCTTACTACATGGGGATCAGCATCAGCATCCTGGAAATCATCAAGCAA  906

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAGTGACCCCTCGCCTTC  660
          |||| |||| || ||| || ||||||| || |||| ||||| ||
Db  907 GGGTGTGAGTTTGAGAACACTGTGCACAGTGATTTCCACACCCAGGGCCCCTAGCTTTC  966

Qy  661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
          ||||||| | |||||||||||||||||
Db  967 TTCCACTGTTGTGTGAACCCCATCCTCTAT  996
```

```
RESULT 14
US-08-802-627A-3
```

Right column

```
; Sequence 3, Application US/08802627A
; Patent No. 5892017
; GENERAL INFORMATION:
;    APPLICANT: Lee, James
;    APPLICANT: Wood, William I.
;    TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
;    NUMBER OF SEQUENCES: 6
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Genentech, Inc.
;       STREET: 460 Point San Bruno Blvd
;       CITY: South San Francisco
;       STATE: California
;       COUNTRY: USA
;       ZIP: 94080
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/802,627A
;       FILING DATE: 19-Feb-1997
;       CLASSIFICATION: 514
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/284586
;       FILING DATE: 10-AUG-1994
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/076093
;       FILING DATE: 11-JUN-1993
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 07/810782
;       FILING DATE: 19-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Love, Richard B.
;       REGISTRATION NUMBER: 34,659
;       REFERENCE/DOCKET NUMBER: P0706P2P1D2
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 415/225-5530
;       TELEFAX: 415/952-9881
;       TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 1737 base pairs
;       TYPE: Nucleic Acid
;       STRANDEDNESS: Single
;       TOPOLOGY: Linear
; US-08-802-627A-3
```

```
Query Match          75.0%;   Score 517.2;   DB 2;   Length 1737;
Best Local Similarity 85.8%;   Pred. No. 6.8e-140;
Matches  592;   Conservative  0;   Mismatches  83;   Indels  15;   Gaps  1;

Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
          |||||||||||||||  ||||||||| | ||||||||||| ||| ||||||||||||||||
Db  322 CTGCACCTGTCAGTGGCCGACCTCCTTCCCTTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  381

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
          |||  ||| ||||||||||||||||||   |  ||  |||||||||| | ||||||||| |
Db  382 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  441

Qy  121 GTCAACCTCTACAGCAGCGGTTCCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
          ||||||||||||||||| || |||||||||||||||||| | |||||||| ||||| |||
Db  442 GTCAACCTCTACAGCAGTGTCCTTCATCCTGGCCTTCATCAGTCTGGACCGCTACCTGGCC  501

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
          || |||||||||||||||||||| | ||||||||||| | | ||||||||||| |||||||
Db  502 ATCGTCCACGCCACCAACAGTCAGGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT  561

Qy  241 GTGGGCGTCTGGATCCCCAGCCCTCCTCCTGACTATACCTGGTCATCTTTGCCGACGTC   300
          || |||||||||||||||  |||||||  | ||| || || ||||||||||||| |||||
Db  562 GTTGGCGTCTGGATCCCCTGCCCTCCTGCTATTCCCGACTTCATCTTTGCCAACG--    619

Qy  301 AGCCAGGGGACATCAGTCAGGGGATGATGACAGTGACATCTGTGACCGGCTTTACCCCGAT  360
```

;   TYPE: Nucleic Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Linear
US-08-801-238-3

Query Match   75.0%;   Score 517.2;   DB 2;   Length 1737;
Best Local Similarity 85.8%;   Pred. No. 6.8e-140;
Matches 592;   Conservative 0;   Mismatches 83;   Indels 15;   Gaps 1;

```
Qy    1 CTGCACCTGTCAGTGGCTGGCCGACCCTCCTCTTTGTCATCACACTCGGGCAGTTGAT 60
          |||||||||||||||||| |  ||||||||||||||||||||||| | || |||||||||
Db  322 CTTGCACCTGTCAGTGGCCGCCGACCCTCCTCTTTGTCATCATCAGCTTCCTTCTGGGCAGTTGAT 381

Qy   61 GCCATGGCTGACTGGTACTTGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
          ||||||| ||||||||||||||||||||||||  |||||||||||||||||||||||||
Db  382 GCCGGTGCCAAACTGGTACTTGGGAACTTCCTATGCAAGCAGTCAGTCCATGTCATCTACACA 441

Qy  121 GTCAACCTTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCGGACCGGTACCTCGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  442 GTCAACCTTACAGCAGCAGTGTTCTCATCCTGGCCTTCATCAGTCGGACCGGTACCTCGGCC 501

Qy  181 ATTGTCCAGCCACCAACAAGTCAAAGGCCAAGGAAAACTCGGCTGAAAAGGCAGTCTAT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  502 ATCGTCCAGCCACCAACAAGTCAGAGGCCAAGGAAGTCTGGCTGAAAAAGGTGGTCTAT 561

Qy  241 GTGGGCGTCTGGGATCCCAGCCCTCCTGCCTCCTGCGCTATTCCCGACTATTCTTGCCGACGTC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  502 GTTGGCGTCTGGGATCCCTGCCCTCCTGCGCTATCCCCGACTATTCATCTTGCCAACG-- 619

Qy  301 AGCCAGGGGACATCAGTCAGGGGATGACGATGACAGATATAATGTGACCGCCTTACCCCGAT 360
          ||||||||||||||||||||||||||
Db  620 ----------TCAGTGAGGCAGATGACAGATATAATCGTGACCGCTTCTACCCCAAT 666

Qy  361 AGCCTGTGGATGGTGCGTGTTCAATTCCAGCATCAGCATATAATGTGGGTCTCATCTGCCGCGGC 420
          ||||||||||||| |||||||||||||||||||||||||||||||||| ||
Db  667 GACTTGTGGGGTGGTTGGTGTTGTTCCAGTTCAGCACATACTCTTCAGCCTTATCCTCGCCCTGGT 726

Qy  421 ATCGTCATCCTCCTCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC 480
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  727 ATTGTCATCCTCCTGTCTCGTATTGCATTATCATCCTCCAAGCTGTCACACTCCAAGGGGCCAC 786

Qy  481 CAGAAGCGCAAGGCCCTCAGACACAGTCACTCCATCCTCATCCTCAGTCTCCTTCTTCTTGCCTGTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  787 CAGAAGCGCAAGGCCCTCAAGACCACAGACCACAGTCATCCTCACTCGGCTTTCTTGGCTTTCTTCGCCTGTTGG 846

Qy  541 CTGCCATATTATGTGGGGATCAGCATCAGCACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  847 CTGCCTTACTACATTGGGATCAGCATCAGCATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAA 906

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGGCCCTCGCCTTC 660
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  907 GGGGTGAGTTTGAGAACACTGTGCACAAGTGTGCACAAGTGGATTTCCATCACCGAGGGCCCTAGCTTTC 966

Qy  661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
          |||||||||||||||||||||||||||||
Db  967 TTCCACTGTTGTCTGAACCCCATCCTCTAT 996
```

RESULT 15
US-08-801-238-3
;   Sequence 3, Application US/08801238
;   Patent No. 5919896
;   GENERAL INFORMATION:
;   APPLICANT: Lee, James
;   APPLICANT: Wood, William I.
;   TITLE OF INVENTION: PF4A RECEPTOR
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/801,238
;   FILING DATE: 19-Feb-1997
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/284586
;   FILING DATE: 10-AUG-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/076093
;   FILING DATE: 11-JUN-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/810782
;   FILING DATE: 19-DEC-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B.
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: P0706P2P1D1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
;   INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1737 base pairs

Search completed: May 17, 2004, 16:15:43
Job time : 59.2446 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 12:11:00 ; Search time 337.746 Seconds
                                 (without alignments)
                                 9271.041 Million cell updates/sec

Title: US-09-367-052-3
Perfect score: 690
Sequence: 1 ctgcacctgtcagtggctga.........gcctgaacccatcctctat 690

Scoring table: IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched: 2947324 seqs, 226902515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : Published Applications_NA:*
   1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
   2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
   3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
   4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
   5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
   6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
   7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
   8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
   9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
  10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
  11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
  12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
  13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
  14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
  15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
  16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
  17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
  18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
  19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	518	75.1	5161	16	US-10-160-401-1	Sequence 1, Appli
2	517.2	75.0	1059	13	US-09-813-651B-84	Sequence 84, Appl
3	517.2	75.0	1059	15	US-10-014-322A-125	Sequence 125, App
4	517.2	75.0	1059	16	US-10-160-401-2	Sequence 2, Appli
5	517.2	75.0	1225	15	US-10-101-510-741	Sequence 741, App
6	517.2	75.0	1225	16	US-10-305-720-1235	Sequence 1235, Ap
7	517.2	75.0	1664	17	US-10-641-643-1213	Sequence 1213, Ap
8	517.2	75.0	1670	9	US-09-880-107-2143	Sequence 2143, Ap
9	517.2	75.0	1670	10	US-09-960-706-636	Sequence 636, App
10	517.2	75.0	1670	16	US-10-372-683-1	Sequence 1, Appli
11	517.2	75.0	1670	16	US-10-440-464-77	Sequence 77, Appl
12	517.2	75.0	1679	13	US-10-211-462-80	Sequence 80, Appl
13	517.2	75.0	1679	13	US-10-181-906-9	Sequence 9, Appli
14	517.2	75.0	1679	13	US-10-342-887-912	Sequence 912, App
15	517.2	75.0	1679	13	US-10-151-274-14	Sequence 14, Appl
16	517.2	75.0	1679	13	US-10-172-118-912	Sequence 912, App
17	517.2	75.0	1679	13	US-10-170-385-332	Sequence 332, App
18	517.2	75.0	1679	15	US-10-225-567A-75	Sequence 75, Appl
19	517.2	75.0	1679	16	US-10-021-660-58	Sequence 58, Appl
20	517.2	75.0	1679	16	US-10-341-434-177	Sequence 177, App
21	517.2	75.0	1711	10	US-09-971-392-20	Sequence 20, Appl
22	517.2	75.0	1711	15	US-10-101-510-459	Sequence 459, App
23	517.2	75.0	1737	9	US-09-104-063-3	Sequence 3, Appli
24	517.2	75.0	1737	13	US-10-666-689-3	Sequence 3, Appli
25	515.6	74.7	1102	9	US-09-870-759-143	Sequence 143, App
26	515.6	74.7	1102	10	US-09-751-708A-143	Sequence 143, App
27	512.4	74.3	1059	11	US-09-826-509-484	Sequence 484, App
28	505.2	73.2	1224	15	US-10-101-510-674	Sequence 674, App
29	466	67.5	1076	15	US-10-017-161-789	Sequence 789, App
30	466	67.5	1076	16	US-10-292-798-687	Sequence 687, App
31	303.2	43.9	421	9	US-09-796-692-2611	Sequence 2611, Ap
32	303.2	43.9	421	15	US-10-040-862-2611	Sequence 2611, Ap
33	303.2	43.9	421	16	US-10-057-475B-2611	Sequence 2611, Ap
34	303.2	43.9	421	16	US-10-154-884B-2611	Sequence 2611, Ap
35	276.4	40.1	1902	9	US-09-953-692-1	Sequence 1, Appli
36	276.4	40.1	1902	9	US-09-951-717-1	Sequence 1, Appli
37	185.6	26.9	507	13	US-10-085-783A-14955	Sequence 14955, A
38	185.6	26.9	507	16	US-10-242-535A-14955	Sequence 14955, A
39	151.4	21.9	1068	15	US-10-237-561-45	Sequence 45, Appl
40	150	21.7	1119	12	US-09-997-722-162	Sequence 162, App
41	150	21.7	1119	15	US-10-251-385-65	Sequence 65, Appl
42	150	21.7	1679	13	US-09-104-063-5	Sequence 5, Appli
43	150	21.7	1679	13	US-10-666-689-5	Sequence 5, Appli
44	150	21.7	2824	12	US-09-997-722-161	Sequence 161, Appl
45	150	21.7	2824	15	US-10-225-567A-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-10-160-401-1
; Sequence 1, Application US/10160401
; Publication No. US20030207281A1
; GENERAL INFORMATION:
; APPLICANT: Genaissance Pharmaceuticals, Inc.
; APPLICANT: Bentivegna, Steven C.
; APPLICANT: Bieglecki, Karyn M.
; APPLICANT: Koshy, Beena
; APPLICANT: Monroe, Glen
; APPLICANT: Rounds, Eileen
; TITLE OF INVENTION: HAPLOTYPES OF THE CXCR4 GENE
; FILE REFERENCE: MWH-0121US
; CURRENT APPLICATION NUMBER: US/10/160,401
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: PCT/US01/12268
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,025
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5161
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3118)..(3118)
; OTHER INFORMATION: PS1: polymorphic base thymine or adenine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3583)..(3583)
; OTHER INFORMATION: PS2: polymorphic base cytosine or thymine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3952)..(3952)
; OTHER INFORMATION: PS3: polymorphic base cytosine or thymine

US-09-813-651B-84

Query Match   75.0%;   Score 517.2;   DB 13;   Length 1059;
Best Local Similarity   85.8%;   Pred. No. 1.2e-156;
Matches 592;   Conservative   0;   Mismatches 83;   Indels 15;   Gaps 1;

```
Qy     1 CTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT    60
         ||||||||||||||||| ||||| ||||  |||||||||| || |||||||||||||||
Db   232 CTGCACCTGTCAGTGGCCGACCTCCTCCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT   291

Qy    61 GCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
         ||| |||| | ||||||||||||| | | |  | | ||||| |||||| |||||||||
Db   292 GCCGTGGCAAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA   351

Qy   121 GTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC   180
         |||||||||||||||| ||| | ||||||||||||||||||| |||||||| |||||| ||
Db   352 GTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCCGTACCTGGCC   411

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
         || |||||||||||||||||||  | |||||||||| |||||||||||||||||| |||||
Db   412 ATCGTCCACGCCACCAACAGTCAGGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT   471

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTCGACTATACCTGACTATTCCCGACTCTTTGCCGACGTC   300
         ||||||||||||||||||||||||| || ||| | |||| ||||||| ||| |||||||||| 
Db   472 GTTGGGCGTCTGGATCCCTGCTCCTCGCTGCTATTCCGGACTATTCCCGACTTCATGCCAACG--   529

Qy   301 AGCCAGGGGGACATCAGTCAGGGGATGACAGTGACAGTTTACCCCGAT   360
                  |||||||||||| ||| |||  ||||||  |||||||||||
Db   530 ------------TCAGTGAGGCGATGACAGATGACAGATATCTGACCGCTTCTACCCCAAT   576

Qy   361 AGCCTGTGATGGTGTTCAATTCAGCATATAATGTGGGTCTCATCCTGCCCGGC   420
         ||||||||| | |||| || || | | || ||| | ||| |||| |||| |||
Db   577 GACTTGTGGGTGGTTGTGTTGCAGTTTCAGCACATCATTGCCTTATCCTGCCTGGT   636

Qy   421 ATCGTCATCCTCTCCTGTTACTGCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC   480
         |||||||||||||||||||||||| | ||||| || || |||||||||| |||||||
Db   637 ATTGTCATCCTGTCCTGCTATTGCATTATCTCTCCAAGCTGTCACACTCCAAGGGCCAC   696

Qy   481 CAGAAGCGCAAGGCCCTCAAGACGACAGTCAGGGGATGACAGTTCTTTCTTTGCCTGCTGG   540
         ||||||||||||||||||||||||| ||||||||| |||||||| |||||||||||||||
Db   697 CAGAAGCGCAAGGCCCTCAAGACCACACCAGTCATCCTCATCCTCCTCTCCTGCCTGTTGG   756

Qy   541 CTGCCATATATGTGGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAA   600
         |||||||| ||| || | ||||  ||||| ||||||||||| |||||||||
Db   757 CTGCCTTACTACATTGGATCAGCATCAGCTCCTCCTTCATCTCCTGGAAATCATCAAGCAA   816

Qy   601 GGATGTGACTTGCGAGAGCATTGTGCACAAGTGACATTCCATCACACAGAGGGCCCTCGCCTTC   660
         |||||||||||| ||||| |||| || || |||||| || |||||||||||||||||||||||
Db   817 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGCAAGTGGGATTTCCATCACCGAGGCCTAGCTTTC   876

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT   690
         |||||||||||||||||||||||||||||
Db   877 TTCCACTGTTGTCTGAACCCCATCCTCTAT   906
```

US-10-160-401-1

Query Match   75.1%;   Score 518;   DB 16;   Length 5161;
Best Local Similarity   85.7%;   Pred. No. 1.4e-156;
Matches 591;   Conservative   2;   Mismatches 82;   Indels 15;   Gaps 1;

```
Qy     1 CTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT    60
         ||||||||||||||||| ||||| ||||  |||||||||| || |||||||||||||||
Db  3401 CTGCACCTGTCAGTGGCCGACCTCCTCCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  3460

Qy    61 GCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
         ||| |||| | ||||||||||||| | | |  | | ||||| |||||| |||||||||
Db  3461 GCCGTGGCAAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  3520

Qy   121 GTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC   180
         |||||||||||||||| ||| | ||||||||||||||||||| |||||||| |||||| ||
Db  3521 GTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCCGTACCTGGCC  3580

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
         || ||||||||||||||||||  | |||||||||| |||||||||||||||||| |||||
Db  3581 ATYGTCCACGCCACCAACAGTCAGGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT  3640

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTCGACTATACCTGACTATTCCCGACTCTTTGCCGACGTC   300
         ||||||||||||||||||||||||| || ||| | |||| ||||||| ||| |||||||||| 
Db  3641 GTTGGGCGTCTGGATCCCTGCTCCTCGCTGCTATTCCGGACTATTCCCGACTTCATGCCAACG--  3698

Qy   301 AGCCAGGGGGACATCAGTCAGGGGATGACAGTGACAGTTTACCCCGAT   360
                  |||||||||||| ||| |||  ||||||  |||||||||||
Db  3699 ------------TCAGTGAGGCGATGACAGATGACAGATATCTGACCGCTTCTACCCCAAT  3745

Qy   361 AGCCTGTGATGGTGTTCAATTCAGCATATAATGTGGGTCTCATCCTGCCCGGC   420
         ||||||||| | |||| || || | | || ||| | ||| |||| |||| |||
Db  3746 GACTTGTGGGTGGTTGTGTTGCAGTTTCAGCACATCATTGCCTTATCCTGCCTGGT  3805

Qy   421 ATCGTCATCCTCTCCTGTTACTGCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC   480
         |||||||||||||||||||||||| | ||||| || || |||||||||| |||||||
Db  3806 ATTGTCATCCTGTCCTGCTATTGCATTATCTCTCCAAGCTGTCACACTCCAAGGGCCAC  3865

Qy   481 CAGAAGCGCAAGGCCCTCAAGACGACAGTCAGGGGATGACAGTTCTTTCTTTGCCTGCTGG   540
         ||||||||||||||||||||||||| ||||||||| |||||||| |||||||||||||||
Db  3866 CAGAAGCGCAAGGCCCTCAAGACCACACCAGTCATCCTCATCCTCCTCTCCTGCCTGTTGG  3925

Qy   541 CTGCCATATATGTGGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAA   600
         |||||||| ||| || | ||||  ||||| ||||||||||| |||||||||
Db  3926 CTGCCTTACTACATTGGATCAGCATCAGCTCCTCCTTCATCTCCTGGAAATCATCAAGCAA  3985

Qy   601 GGATGTGACTTGCGAGAGCATTGTGCACAAGTGACATTCCATCACACAGAGGGCCCTCGCCTTC   660
         |||||||||||| ||||| |||| || || |||||| || |||||||||||||||||||||||
Db  3986 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGCAAGTGGGATTTCCATCACCGAGGCCTAGCTTTC  4045

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT   690
         |||||||||||||||||||||||||||||
Db  4046 TTCCACTGTTGTCTGAACCCCATCCTCTAT  4075
```

RESULT 2
US-09-813-651B-84

RESULT 3
US-10-014-322A-125
; Sequence 125, Application US/10014322A
; Publication No. US20030167129A1
GENERAL INFORMATION:
; APPLICANT: Nestor, Jr., John
; APPLICANT: Wilson, Carol
; APPLICANT: Tan Hehir, Christina
; APPLICANT: Kates, Steven
; APPLICANT: Krstenansky, John
; TITLE OF INVENTION: Binding Compounds and Methods for Identifying Binding Compounds
; FILE REFERENCE: CNS-008
; CURRENT APPLICATION NUMBER: US/10/014,322A
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/243,587
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 09/813,651
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 09/813,653
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 09/813,448
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 125
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1059)
; OTHER INFORMATION: human CXCR4 gene
US-10-014-322A-125

Query Match   75.0%;   Score 517.2;   DB 15;   Length 1059;
Best Local Similarity 85.8%;   Pred. No. 1.2e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCCTTCTTGTCATCACACTCCCCTTCTTGGGCAGTTGAT   60
        |||||||||||||||| | ||||| | | ||||||||||| ||| |||| ||||||||||||
Db  232 CTGCACCTGTCAGTGCCGACCTCCTCTCTTGTCATCACGCTTCCCTTCCTTGGGCAGTTGAT  291

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTGTGTAAGGCTGTCCATATCATCTACACT     120
        ||| ||||  |||||||||||||||||  | | || ||||| |||||| |||||||||
Db  292 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA     351

Qy  121 GTCAACCTCTACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC   180
        ||||||||||||| || | |  ||||||||||||||||||||||||||||||||||| |||
Db  352 GTCAACCTCTACAGCAGCAGTCGTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCC   411

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
        |||||||  |||| || |||||||||||||||||||| | | ||||| ||||| |||||||
Db  412 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGCTGGCTGAAAAGGTGGTCTAT   471

Qy  241 GTGGGGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGCACGTC     300
        | |||| |||||||  ||||| |||| ||||||||| |||||||||||||||  |||
Db  472 GTTGGGCCTGGATCCCGCCCTCCTGCTCTGACTATTTCCCGACTTCATCTTTGCCCAACG--  529

Qy  301 AGCCAGGGGGACATCAGGTCAGGGGGATGACAGAGTACATCTGTGACCGCCTTCACCCCGAT   360
        |||||||| |||||||||||||||| | |||| | ||||||| || |||||| || ||| |
Db  530 ------------TCAGTGAGGCAGATGAGATGACAGATATCTGTGACCGCCTTTACCCCAAT   576

Qy  361 AGCCTGGTGATGGTGGTTGTTCCAGTTTCCAGTTTCAGCAGCATGATAATGGTGGTCTCACGTC   420
        ||||||||||||| ||||||||||||||||||||||  ||||| || | |||| |||||
Db  577 AGCCTGGTGATGGTGGTTGTTCCAGTTTCCAGTTTCAGCAGCATCATGATAATGGTGGTCTCACGTC   636

Qy  421 ATCGTCATCCTCTCCGTGTTACTGCATCATCATCTCTAAGTCGTCACACTCCAAGGGCCAC   480
        |||||||||||||||  || | |||||||||||||||  ||| |||||||||||||||||
Db  637 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGTCGTCACACTCCAAGGGCCAC   696

Qy  481 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTGCTTCTTCGGCTTGTTGG   540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    —
```

RESULT 4
US-10-160-401-2
; Sequence 2, Application US/10160401
; Publication No. US20030207281A1
GENERAL INFORMATION:
; APPLICANT: Genaissance Pharmaceuticals, Inc.
; APPLICANT: Bentivegna, Steven C.
; APPLICANT: Bieglecki, Karyn M.
; APPLICANT: Koshy, Beena
; APPLICANT: Monroe, Glen
; APPLICANT: Rounds, Eileen
; TITLE OF INVENTION: HAPLOTYPES OF THE CXCR4 GENE
; FILE REFERENCE: MWH-0121US
; CURRENT APPLICATION NUMBER: US/10/160,401
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: PCT/US01/12268
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,025
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-160-401-2

Query Match   75.0%;   Score 517.2;   DB 16;   Length 1059;
Best Local Similarity 85.8%;   Pred. No. 1.2e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

```
Db  697 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTGCTTCTTCGGCTTGTTGG   756

Qy    1 CTGCACCTGTCAGTGGCTGACCTCCCTTCTTGTCATCACACTCCCCTTCTTGGGCAGTTGAT   60
        |||||||||||||||| | ||||| | | ||||||||||| ||| |||| ||||||||||||
Db  232 CTGCACCTGTCAGTGCCGACCTCCTCTCTTGTCATCACGCTTCCCTTCCTTGGGCAGTTGAT  291

Qy  541 CTGCCATATTATGTGGGATCAGCATCGACTCCTTCCTCCTTTGGGACTCATCAAGCAA   600
        ||| |||| || ||||||||||||| ||||||||| || ||| |||   |||||||||
Db  757 CTGCCTTACTACATTGGGATCAGCATCGACTGACTCCTTCATCCTCCTGGAAATCATCAAGCAA  816

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTGTGTAAGGCTGTCCATATCATCTACACT   120
        ||| ||||  ||||||||||||||||  | | || ||||| |||||| |||||||||
Db  292 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  351

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGACTCCATCACAGAGCCCTCGCCTTC   660
        || ||||| ||||||| | || || |||||||| | ||||||| || ||| || |||
Db  817 GGGTGTGAGTTTGAGAACACTGTCCACAAGTGATTTCCATCACCGAGGCCCAGCTTTC  876

Qy  121 GTCAACCTCTACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC   180
        ||||||||||||| || | |  ||||||||||||||||||||||||||||||||||| |||
Db  352 GTCAACCTCTACAGCAGCAGTCGTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCC   411

Qy  661 TTCCACTGTGCCTGAACCCATCCTCTAT   690
        |||||||||  | |||||| |||||||||
Db  877 TTCCACTGTTGTGTCTGAACCCCATCCTCTAT  906

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
        |||||||  |||| || |||||||||||||||||||| | | ||||| ||||| |||||||
Db  412 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGCTGGCTGAAAAGGTGGTCTAT   471

Qy  241 GTGGGGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGCACGTC     300
        | |||| |||||||  ||||| |||| ||||||||| |||||||||||||||  |||
Db  472 GTTGGGCCTGGATCCCGCCCTCCTGCTCTGACTATTTCCCGACTTCATCTTTGCCCAACG--  529

Qy  301 AGCCAGGGGGACATCAGGTCAGGGGGATGACAGAGTACATCTGTGACCGCCTTCACCCCGAT   360
        |||||||| |||||||||||||||| | |||| | ||||||| || |||||| || ||| |
Db  530 ------------TCAGTGAGGCAGATGAGATGACAGATATCTGTGACCGCCTTTACCCCAAT   576

Qy  361 AGCCTGGTGATGGTGGTTGTTCCAGTTTCCAGTTTCAGCAGCATGATAATGGTGGTCTCACGTC   420
        ||||||||||||| ||||||||||||||||||||||  ||||| || | |||| |||||
Db  577 GACTTGTGGGTGGTTGTGTTCCAGTTTGTGTTCCAGTTTCAGCACATCATGATAATCCTGCCTGGT   636
```

Db   602 GACTTGTGGGTGGTTGTGTTCCAGTTTCAGCACACATCATGGTTGGCCTTATCCTGCCTGGT 661
Qy   421 ATCGTCATCCTCCTGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCAC 480
          |||||||||||| || |||| |||| ||||| |||||||||||||||||||||
Db   662 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 721

Qy   481 CAGAAGCGCAAGGCCCTCAGAGCGACAGTCATCCTCAGCTTCTTGCCTGTGG 540
          ||||||||||||||||||||||| ||||||||||| |||| || | |||||||
Db   722 CAGAAGCGCAAGGCCCTCAGACCACAGTCATCCTCGGCTTTCTTCTTCGCCTGTGG 781

Qy   541 CTGCCATTATGTGGGATCAGCATGCATCGACTCCTTCATCTTTGGGAGTCATCAAGCAA 600
          ||||| | || ||||||||||| |||||||||||||||||| |||| || ||||||||
Db   782 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCTTCATCTTCCTGGAAATCATCAAGCAA 841

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
          || ||||| ||||| | | ||||||||||||| ||||||||| ||||||||| || |||
Db   842 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 901

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
          |||||||||||| ||||||||||||||||||
Db   902 TTCCACTGTTGTCTGAACCCCATCCTCTAT 931

RESULT 6
US-10-305-720-1235
; Sequence 1235, Application US/10305720
; Publication No. US2004001013GA1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1235
; LENGTH: 1225
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: GenBank ID No. US2004001013GA1 g189313
US-10-305-720-1235

Query Match   75.0%; Score 517.2; DB 16; Pred. No. 1.3e-156; Length 1225;
Best Local Similarity   85.8%;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
          |||||||||||||||||||||||||||||| |||||||| | | ||||||||||||||
Db   257 CTGCACCTGTCAGTGGCCGACCTCCTTTGTCATCACACGCTTCCCTTCTGGGCAGTTGAT 316

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
          || ||||||||||| |||||||||||||||||| |||| |||||||||||||| | |||
Db   317 GCCGTGGCAAACTGTACTTTGGGAACTTCCTATGCAAGGCAGTTCCATGTCATCTCACACA 376

Qy   121 GTCAACCTTCACAGCAGCGGTTCTCATCAGCCTGGCCTTCATCAGCCTGACCTGCCGCC 180
          ||||||||||||||| ||| || || ||||||||||||||||||||| ||||||||||
Db   377 GTCAACCTTCACAGCAGTGTCTCCTCATCAGCCTGCCTTCATCAGCCTGACCGGCCTACCTGGCC 436

Qy   181 ATTGTCCACGCCACCACAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
          ||||||||||||||||||||||||||| |||||||||||| | || ||||||||||| ||||
Db   437 ATCGTCCACGCCACCACAACAGTCAGAGGCCAAGGAAGCAGCGTGTTGGCTGAAAAGGTGGTCTAT 496

Qy   241 GTGGGCGTCTGGATCCCCAGCCCTCCTCCTGACTTCATCTTTGCCGACGTC 300
          ||||| |||||||||||| |||||||| |||| ||| |||||||||| ||||
Db   497 GTTGGCGTCTGGATCCCGCCTCCTGCTATTCCCGACTATTCATCTTTGCCAACG-- 554

Qy   301 AGCCAGGGGACATCAGTGGGATGCACATCTGTGACCGCCTTTACCCCGAT 360
          ||||||||| |||||| ||||||| |||| ||||| || | ||||
Db   555 ------------TCAGTGAGGCAAACTGGTTACTTTGGAACTTCCATGCAACCG-- 601

Qy   361 AGCCTGTGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGGC 420
          ||| || | |||||||| || |||||||||||| ||||||||||||||||
Db   555 TCAGTGAGGCAGATGACAGATCAGGGATGACAGAGGTACATCTGTGACCGCTTACCCCGAT 601

US-10-641-643-1213

Query Match   75.0%;   Score 517.2;   DB 17;   Length 1664;
Best Local Similarity   85.8%;   Pred. No. 1.5e-156;
Matches 592;   Conservative   0;   Mismatches   83;   Indels   15;   Gaps   1;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCTTCGGGCAGTTGAT   60
        |||||||||||||||| |||||||||| ||||||||||| || ||||||||||||||
Db  308 CTGCACCTGTCAGTGGCGCCGACCTCCTCTTGTCATCACGCTTCCCTTCGGGCAGTTGAT  367

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTAAGGCTGTCCATATCATCACACT  120
        ||| |||||  ||| ||||||||||||  ||  | |||||||||  || |||| ||||
Db  368 GCCGTGGCAAACTGTACTTTGGGAACTTCCTATGCAAGGCAGTGCATCATCACACA  427

Qy  121 GTCAACCTTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACTTCGC  180
        |||||||| |||||||||  | ||||| || ||||||||| |||||||| ||| ||||
Db  428 GTCAACCTCTACAGCAGCTGTCCTCATCAGTCCTCATCAGTCTGGACCTACCTGGCC  487

Qy  181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
        || |||||| ||||||||||||| ||| |||||   |||| |||||||| ||| ||||||
Db  488 ATCGTCCAGCCACCAACAGTCAGAGGAGCGTGTTGGCTGAAAAAGGTGGTCTAT  547

Qy  241 GTGGGCGTCTGGATCCCAGCCCTCCTCGACTATACCGACTTCATCTTTGCCGACGTC  300
        || |||||||||||||| ||| ||||| ||||||  ||||||||||||||| |||| 
Db  548 GTTGGCGTCTGGATCCCTGCCTCCTGCGACTATTCCCGACTTCATCTTTGCCAACG--  605

Qy  301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGCAGGTACATCGTGACACGCCCTTACCCCGAT  360
        ||||| |||| ||||  | || ||||| ||||||| |  |||||||| || || |||||||
Db  606 ----------TCAGTGAGGCAGATGACAGCAGATATATCGTGACCGCGCTTCTACCCCAAT  652

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC  420
        |||| ||||| ||||||||| |||||||| |||| ||||| ||| |||||||||||||
Db  653 GACTTGTGGGTGGGTGGTGTTGTTCCAGTTTCAGCACATCATGGTTGGCCTTATCCTGCCTGGT  712

Qy  421 ATCGTCATCCTCCGTACTGCATCATCATCTCAAGCTGTCACACTCCCAAGGGCCAC  480
        || ||||||||  | |||| |||||||||| |||||||||||||||||||||||||
Db  713 ATTGTCATCCTGTCCTATTGCATTATCATCTCCAAGCTGTCACACTCCCAAGGGCCAC  772

Qy  481 CAGAAGCGCAAGGCCCTCAAGACGACAGTGCATCCTCATCCAGCTTTCTTGCCTGCTGG  540
        |||||||||||||||||||||||| ||||| |||||||| ||| ||||| |||||| |||
Db  773 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTCGGCTTTCTCGCCTGTTGG  832

Qy  541 CTGCCATATTATGTGGGGATCAGCATGCAGCATGGCATCCTTCCTTTTGGGAGTCATCAAGCAA  600
        ||||| || | | ||||||||||||||||||||||||||  ||||| ||||  ||||||||||||
Db  833 CTGCCTTACACATTGGGATCAGCATGCAGCATGGCATCCTTCATCCTCTGGAAATCATCAAGCAA  892

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTGCACAAGTGGATCCCTCGCCTTC  660
        || |||||| || ||| |||||||||||||||| |   | ||||||| |||| ||||||
Db  893 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGGATTTCCATCACCGAGCCCTAGCTTTC  952

Qy  661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
        |||||||||||| |||||||||||||||||
Db  953 TTCCACTGTTGTCTGAACCCCATCCTCTAT  982
```

RESULT 8
US-09-880-107-2143
; Sequence 2143, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950

```
Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC  420
        |||| ||||| ||||||||| |||||||| |||| ||||| ||| |||||||||||
Db  602 GACTTGTGGGTGGGTGGTGTTCAGTTCCAGTTTCAGCACATCATGGTTGGCCTTATCCTGCCCTGGT  661

Qy  421 ATCGTCATCCTCCGTACTGCATCATCATCTCAAGCTGTCACACTCCCAAGGGCCAC  480
        || |||||||| | ||| |||||||||| |||||||||||||||||||||||||||
Db  662 ATTGTCATCCTGCCTATTGCATTATCATCTCCAAGCTGTCACACTCCCAAGGGCCAC  721

Qy  481 CAGAAGCGCAAGGCCCTCAAGACGACAGTGCATCAGCATGGATCCTTCATCCTTTGGGAGTCATCAAGCAA  540
        |||||||||||||||||||||||| ||||| |  | | ||||| || ||||||| ||| |||||||||||
Db  722 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCATCAGTCCTTCATCCTCTGGCTTCTTCGCCTGTTGG  781

Qy  541 CTGCCATATTATGTGGGGATCAGCATGCATGCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA  600
        ||||| || | | ||||||||||||||||| ||||||||||||| |||  ||||||||||||
Db  782 CTGCCTTACACATTGGGATCAGCATGCGACTCCTTCATCCTCTGGAAATCATCAAGCAA  841

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCCTTCATCACAGAGGCCCTCGCCTTC  660
        || |||||| || ||| |||||||||||||||| | | | ||||||| ||||| ||||||
Db  842 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGGATTTCCATCACCGAGCCCTAGCTTTC  901

Qy  661 TTCCACTGTTGCCTGAACCCCATCCTCTAT  690
        |||||||||||| |||||||||||||||||
Db  902 TTCCACTGTTGTCTGAACCCCATCCTCTAT  931
```

RESULT 7
US-10-641-643-1213
; Sequence 1213, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
;   Susan G. Stuart
;   Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
;   GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/641,643
;   FILING DATE: 14-Aug-2003
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: <Unknown>
;   FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1213:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1664 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: 9219868
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1213:

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2143
; LENGTH: 1670
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L06797
US-09-880-107-2143

Query Match   75.0%;   Score 517.2; DB 9; Length 1670;
Best Local Similarity 85.8%;   Pred. No. 1.5e-156;
Matches   592; Conservative   0; Mismatches 83; Indels 15; Gaps   1;

```
Qy      1  CTGCACCTGTCAGTGGCTGACCCCTCTTTGTCATCACACTCCCTTCTGGGCAGTTGAT  60
           ||||||||||||||||||||||||| ||| |||| ||||||||||||||||||||||||
Db    304  CTGCACCTGTCAGTGGCCGGACCTCCTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  363

Qy     61  GCCATGGCTGACTGGTACTTTGGGAAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
           ||||||| ||||| ||||||||||| | |||||||||  |||||| ||||||||||| ||
Db    364  GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCATCTCATCTACACA  423

Qy    121  GTCAACCTCACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    424  GTCAACCTCACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC  483

Qy    181  ATTGTCCAGCGCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
           ||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    484  ATTGTCCAGCGCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  543

Qy    241  GTGGGCGTCTGGATCCCAGCCTCCTGACTATACCGACCGGCCTTTGCCGACGTC  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    544  GTTGGCGTCTGGATCCCTGCCTCCTGCTTCATCTTTCCCGACTTCATCTTTGCCAACG--  601

Qy    301  AGCCAGGGGGACATCAGTCAGGGGATGACAGTACATCGTGACCGCCTTTACCCCGAT  360
           |||||||||||||||||||||||||||||||||||||||||||||
Db    602  ----------TCAGTGAGGCAGATGACAGATATATCGTGACGACTTCATCTTTCACCCCAAT  648

Qy    361  AGCCTGTGGATGGTGGTGTTCAATTCCAGCATCATCATTCAATTCCAGCATCATCATCCAAGGGCCAC  420
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    649  GACTTGTGGGTGGTTGTGTTCGTTCCAGTTTCAGCACACATCATGGTTGGCCTTATCCTGCCTGGT  708

Qy    421  ATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCTAAGCTGTCAAGGGCCAC  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709  ATTGTCATCCTGTTCCTGCTATTGCATTATCATCTCCAAGCTGTCACAACTCCAAGGGGCAC  768

Qy    481  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTTCCTCCTTTTGGGAGTCATCAAGCAA  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769  CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCTACCTGCTTTCTTCGCCTGTTGG  828

Qy    541  CTGCCATATTATGTGGGAGCATTGTGCACAGTGGATCCTCATCACAGAGGCCCTGCCTTC  600
           ||||| |||||||||||||||||||| ||| ||||||||||||||||||||||||||||
Db    829  CTGCCTTACTACATTGGGATCAGCATGGGCATCGACTCCTCTGAAATCATCAAGCAA  888

Qy    601  GGATGTGACTTCGAGAGCATTCGAGAGCATTGAGAACACTGAGTTCATCAACCCCATCCTCTAT  660
           ||||||||||||||||||||||||||||||||
Db    889  GGGTGTGAGTTTGAGAACACTGAGTTCATCAACACCGAGTTCCATCACCGAGCCCTAGCTTTC  948

Qy    661  TTCCACTGTTGCCTGAACCCATCCTCTAT  690
           |||||||||||||||||||||||||||||
Db    949  TTCCACTGTTGTCTGAACCCCATCCTCTAT  978
```

RESULT 9
US-09-960-706-636
; Sequence 636, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasi
; TITLE OF INVENTION: Gene Expression Profiles
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 636
; LENGTH: 1670
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 L06797
US-09-960-706-636

Query Match   75.0%;   Score 517.2; DB 10; Length 1670;
Best Local Similarity 85.8%;   Pred. No. 1.5e-156;
Matches   592; Conservative   0; Mismatches 83; Indels 15; Gaps   1;

```
Qy      1  CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCTTCTGGGCAGTTGAT  60
           |||||||||||||||||| |||||||| ||||||||||| ||| ||||||||||||||||
Db    304  CTGCACCTGTCAGTGGCCGGACCTCCTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  363

Qy     61  GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT  120
           ||| |||||||||| |||||||||||| ||||| |||||||| ||||||||||||||||
Db    364  GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCATCTCATCTACACA  423

Qy    121  GTCAACCTCACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGCTACCTCCCC  180
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||| |||
Db    424  GTCAACCTCACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGCTACCTGGCC  483

Qy    181  ATTGTCCAGCGCACCAACAGTCTGGCTGGCTGAAAAGGCAGTCTAT  240
           |||||||||||||||||||| |||||||||||||||||||||||||
Db    484  ATCGTCCACGCCACCAACAGTCAGAGGCCTGTGGCTGAAAAGGTGGTCTAT  543

Qy    241  GTGGGCGTCTGGATCCCAGCCTCCTGACTATACCGACCGGCCTTTGCCGACGTC  300
           ||||||||||||||||| ||||||||||| |||| |||| ||| || || 
Db    544  GTTGGCGTCTGGATCCCTGCCTCCTGCTTCATTTCCCGACTTCATCTTTGCCAACG--  601

Qy    301  AGCCAGGGGGACATCAGTCAGGGGGATGACAGTACAGTACATCGTGACCGCCTTTACCCCGAT  360
           |||||||||||||||||||||||||||
Db    602  ----------TCAGTGAGGCAGATGACAGATATATCGTGACGACTTCATCTTTCACCCCAAT  648

Qy    361  AGCCTGTGGATGGTGGTGTTCAATTCCAGCATCATCATTCCAGCATCATCATCCAAGGGCCAC  420
           |||||||| |||||| ||||| |||||||||||||||||||||||||||||||||
Db    649  GACTTGTGGGTGGTTGTGTTCGTTCCAGTTTCAGCACACATCATGGTTGGCCTTATCCTGCCTGGT  708

Qy    421  ATCGTCATCCTCTGTTACTGCATCATCATCTCTAAGCTGTCTAAGCTGTCCAAGGGGCCAC  480
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    709  ATTGTCATCCTGTTCCTGCTATTGCATTATCATCTCCAAGCTGTCACAACTCCAAGGGGCAC  768

Qy    481  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCAAGCAAGCCTTCATCCTTCCTTTTGGGAGTCATCAAGCAA  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769  CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCTACCTGCTTTCTTCGCCTGTTGG  828

Qy    541  CTGCCATATTATGTGGGAGCATTGTGCACAGTGGATCTCCATCACAGAGGCCCTGCCTTC  600
           |||||| ||||| || |||||||||| || |||| ||||||||||||||||||||||||
Db    829  CTGCCTTACTACACTGGGATCAGCATGGGCATCGACTCCTCTGAAATCATCAAGCAA  888

Qy    601  GGATGTGACTTCGAGAGCATTGAGAACACTGAGTTCATCAACCCCATCCTCTAT  690
           ||||||| ||||||||| |||||||||||||
Db    889  GGGTGTGAGTTTGAGAACACTGAGTTCATCAACACCGAGTTCCATCACCGAGCCCTAGCTTTC  948

Qy    661  TTCCACTGTTGCCTGAACCCATCCTCTAT  690
           |||||||||||||||||||||||||||||
Db    949  TTCCACTGTGTGTCTGAACCCCATCCTCTAT  978
```

RESULT 10
US-10-372-683-1
; Sequence 1, Application US/10372683
; Publication No. US20040009171A1
; GENERAL INFORMATION:
; APPLICANT: GERRITSEN, MARY E.

;   APPLICANT: PEALE JR., FRANKLIN V.
;   APPLICANT: WU, THOMAS D.
;   TITLE OF INVENTION: METHODS FOR THE TREATMENT OF CARCINOMA
;   FILE REFERENCE: P1928R1P1
;   CURRENT APPLICATION NUMBER: US/10/372,683
;   CURRENT FILING DATE: 2003-02-21
;   PRIOR APPLICATION NUMBER: US 10/271,690
;   PRIOR FILING DATE: 2002-10-16
;   PRIOR APPLICATION NUMBER: US 60/344,534
;   PRIOR FILING DATE: 2001-10-18
;   NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 1
;   LENGTH: 1670
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-372-683-1

Query Match   75.0%;   Score 517.2;   DB 16;   Length 1670;
Best Local Similarity   85.8%;   Pred. No. 1.5e-156;
Matches   592;   Conservative   0;   Mismatches   83;   Indels   15;   Gaps   1;

```
Qy     1  CTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT    60
          |||||||||||||||||  |||||||| |||||||||| || ||||||||||||||||
Db   304  CTGCACCTGTCAGTGGCCGACCTCCTCCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  363

Qy    61  GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
          ||| |||| ||||||||||||||||||  ||  | ||||| ||||||| ||||||||| |
Db   364  GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  423

Qy   121  GTCAACCTTCACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC   180
          ||||||||  |||||||| || ||||||||||||||||| ||||||||||||||| |||
Db   424  GTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGTACCTCGGCC  483

Qy   181  ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
          || ||||||||||||||||||| ||||||||||||| |||| ||||||||||| |||||||
Db   484  ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT  543

Qy   241  GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC   300
          || ||||||||||||| ||||||||| || ||| ||||||||||||||||||||  |
Db   544  GTTGGGCGTCTGGATCCTGCCCTCCTGCTGCTGCTATTCCGACTTCATCTTTGCCAACG--  601

Qy   301  AGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCCGGATACCCCGAT   360
                      |||| ||| ||| ||| ||||| || || ||||||||  | |||||| |
Db   602  ------------TCAGTGAGGCAGATGACAGATATCTGTGACCCGGTTCTACCCCAAT  648

Qy   361  AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGGGTCTCATCCTGCCGGC   420
          ||| ||||| || ||||| |||| | |||| | ||| | | |||||||||||| ||
Db   649  GACTTGTGGGTGGTGTGTTCAGTTTCAGCTTCAGCATCATGGTTGGCCTTATCCTGCCTGGT  708

Qy   421  ATCGTCATCCTCCTCGTTACTGCATCATCTCTAAGCTGTCACACTCCAAGGGCCAC   480
          || |||||||| ||| | | |||| | |||||| ||||||||||||||||||||||
Db   709  ATTGTCATCCTGCTCGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC  768

Qy   481  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCGACTTCTTTGCCTGCTGG   540
          |||||||||||||||||| | || ||||||||| | |||| ||||| |||
Db   769  CAGAAGCGCAAGGCCCTCAGACCACAGTCATCCTCATCCTGGCTTTCTCGCCTGTTGG  828

Qy   541  CTGCCATATATGTCGGGATCGGGATCAGACCCCATCCTCTAT   600
          ||||| || |  | ||||||| | |||| |||||||||||||
Db   829  CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTCCTGAAATCATCAAGCAA  888

Qy   601  GGATGTGACTTCGAGAGCATTGTGCACAAGTGACATCCGATCTCCATCACAGAGAGGGCCCTCGCCTTC   660
          || |||||| || ||| |  ||||||||||| | || | ||| ||||||| | ||||||||| ||| ||
Db   889  GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGGAGGGCCCCTAGCTTTC  948

Qy   661  TTCCACTGTTGCCTGACCCCATCCTCTAT   690
          |||||||||||  ||| |||||||||||||
Db   949  TTCCACTGTTGTCTGAACCCCATCCTCTAT  978
```

RESULT 11
US-10-440-464-77
; Sequence 77, Application US/10440464

Qy   601 GGATGTGACTTGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   889 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 948

Qy   661 TTCCACTGTTGCCTGAACCCATCCTCTAT 690
          ||||||||| |||||| |||||||||
Db   949 TTCCACTGTTGTCTGAACCCCATCCTCTAT 978

RESULT 12
US-10-211-462-80

Query Match   75.0%;   Score 517.2;   DB 13;   Length 1679;
Best Local Similarity   85.8%;   Pred. No. 1.5e-156;
Matches   592;   Conservative   0;   Mismatches   83;   Indels   15;   Gaps   1;

Qy   1 CTGCACCTGTCAGTGGCTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTCGGCAGTTGAT 60
        |||||||||||||||| ||||||| ||||||||||||||||| |||||||||||||||||
Db   320 CTGCACCTGTCAGTGGCCGACCTCCTTTCCCTTTGTCATCACGCGTTCCCTTCTCGGCAGTTGAT 379

Qy   61 GCCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
       |||||| |||| ||||||||||| | ||||||||||||| ||| || ||||||
Db   380 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCGTCATCTCATCTACACA 439

Qy   121 GTCAACCTCTACAGGCAGCAGTGTCCTCATCCTGGCCTTCATCAGCCGGTACCTCGCC 180
      ||||||||||||||||||| || ||||||||||||| |||||||||||||||||||
Db   440 GTCAACCTCTACAGCAGCAGTGTCCTCATCCTGGCCTTCATCAGTCGGACCGCTACCTCGGCC 499

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGCCAGTCTAT 240
      ||||||||||||| |||| ||||||||| || ||||||||||||||| |||| ||||||
Db   500 ATCGTCCACGCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT 559

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
      ||| ||||||||||||||||| ||||| |||| || ||| | ||||||||||
Db   560 GTTGGGCGTCTGGATCCCGCCCTCCTGCTGCACTCCCGACTTCATCTTTGCCAACG-- 617

Qy   301 AGCCAGGGGGGACATCAGTCAGGGGATGACAGTGGGTTCTCATCCTGCCCGGC 360
      |||| ||||| |||||| ||||||||||||||
Db   618 -------------TCAGTGAGGCAGATGAACAGATATCTGTGACCGGCTTCACCCCAAT 664

Qy   361 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGTGGGTCTCATCCTGCCCGGC 420
      |||| ||||||||||||||||| |||||||||| ||| ||||||||| |||||||| |
Db   665 GACTTGTGGGTGGTTGTGTTCAGTTCAGACACATCATGTTGCCTTATCCTGACCTGT 724

Qy   421 ATCGTCATCCTCTCGTTACTGACATCATCATCCTAAGCTGTGTCACACTCCCAAGGGCCAC 480

Db   725 ATTGTCATCCTGTCCTGCTATTGCATTGCATTATCATCTCCAAGTGTGTCACTCTTCTTGCCTGCTGG 784
      |||||||| || ||| |||||| |||||| |||||| ||||||

Qy   481 CAGAAGCGCAAGGCCCTCAAGACAGACAGTCATCATCTCCAAGTGTGTCACTCTTCTTGCCTGTGG 540
      ||||||||||||||| |||| || |||| |||| ||||||||| || ||||
Db   785 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCCTGGCTTTCTTCGCGTGTTGG 844

Qy   541 CTGCCATATTATGTGGGGATCAGCATGCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA 600
      |||||||||||||||| ||||||||| |||| ||||||||| ||||||||||| |||||
Db   845 CTGCCTTACTACATACATTGGGATCGACATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAA 904

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
      |||||||| ||| ||||| |||||||||||||||| |||||||||||||||||
Db   905 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 964

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
      ||||||||||| |||||||||||||
Db   965 TTCCACTGTTGTCTGAACCCCATCCTCTAT 994

RESULT 13
US-10-181-906-9

Query Match   75.0%;   Score 517.2;   DB 13;   Length 1679;
Best Local Similarity   85.8%;   Pred. No. 1.5e-156;
Matches   592;   Conservative   0;   Mismatches   83;   Indels   15;   Gaps   1;

Qy   1 CTGCACCTGTCAGTGGCTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTCGGCAGTTGAT 60
       |||||||||||||||| ||||||| ||||||||||||||||| |||||||||||||||||
Db   320 CTGCACCTGTCAGTGGCCGACCTCCTTTCCCTTTGTCATCACACGCGTTCCCTTCTCGGCAGTTGAT 379

Qy   61 GCCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
      |||||| |||| ||||||||||| | ||||||||||||| ||| || ||||||
Db   380 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCGTCATCTCATCTACACA 439

Qy   121 GTCAACCTCTACAGGCAGCAGTGTCCTCATCCTGGCCTTCATCAGCCGGTACCTCGCC 180
      ||||||||||||||||||| || ||||||||||||| |||||||||||||||||||
Db   440 GTCAACCTCTACAGCAGCAGTGTCCTCATCCTGGCCTTCATCAGTCGGACCGCTACCTCGGCC 499

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGCCAGTCTAT 240
      ||||||||||||| |||| ||||||||| || ||||||||||||||| |||| ||||||
Db   500 ATCGTCCACGCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTAT 559

Qy   241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
      ||| ||||||||||||||||| ||||| |||| || ||| | ||||||||||
Db   560 GTTGGGCGTCTGGATCCCGCCCTCCTGCTGCACTCCCGACTTCATCTTTGCCAACG-- 617

Qy   301 AGCCAGGGGGGACATCAGTCAGGGGATGACAGTGGGTACATCCTGCCCGGC 360
      |||| ||||| |||||| ||||||||||||||
Db   618 -------------TCAGTGAGGCAGATGAACAGATATCTGTGACCGGCTTCACCCCAAT 664

Qy   361 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGTGGGTCTCATCCTGCCCGGC 420

Db   665 GACTTGTGGGTGGTTGTGTTCCAGTTCAGCACATCATGTTGGCCTTATCCTGCCTGGT 724
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   421 ATCGTCATCCTCCTGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   725 ATTGTCATCCTGTCCGTTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   481 CAGAAGCGCAAGGCCCTCAAGACAGCAGTCATCCTCATCCTCATCCTTTTGCCTGCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   785 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTTTTCTTCGCCTGTTGG 844
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   541 CTGCCATATTATGTGGGGATGGGGATCAGCATCGACTGACTCCTTTTGGGGAGTCATCAAGCAA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   845 CTGCCTTACTACATTGGGGATCAGCATCGACTCCTTCATCCTTCATCCTCCTCGAAATCATCAAGCAA 904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   905 GGGTGTGAGTTTGAGAACACGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 964
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
          ||||||||||||||||||||||||||||||

Db   965 TTCCACTGTTGTCTGAACCCCATCCTCTAT 994

RESULT 14
US-10-342-887-912
; Sequence 912, Application US/10342887
; Publication No. US2004005834A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-912

Query Match   75.0%; Score 517.2; DB 13; Length 1679;
Best Local Similarity 85.8%; Pred. No. 1.5e-156;
Matches   592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
         |||||||||||||||||||||||||||||||||||||| || |||||||||||||||||

Db 320 CTGCACCTGTCAGTGGCCGACCTCCCTTGTCATCACACGCTTCCCTTCTGGGCAGTTGAT 379

Qy   61 GCCATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||

Db 380 GCCGTGGCAAACTGGTACTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA 439

Qy 121 GTCAACCTCCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCTGCC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 440 GTCAACCTTCTACAGCAGCGTCCTCATCCTGGCCTTCATCAGCCTGACCGGCTACCTGCC 499

Qy 181 ATTGTCCACCCCACCACCACAAGTCAAAGGCCAAGAGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 500 ATCGTCCACGCCACCGCCACAAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAAGGTGGTCTAT 559

_____

Db   241 GTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTGCCGACGTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   560 GTTCGGCGTCTGGATCCCTCCGCCCTCCTGCTGCGACTTCCCGACTTCATCTTGCCCAACG-- 617

Db   301 AGCCAGGGGACATCAGTCAGGGGGATGACAGTACATCGTGACCGGCCTTCACCCCGAT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   618 ------------TCAGTGAGGCAGCAGCAGATATATCGTGACCGGCCTTCACCCCCAT 664

Db   361 AGCCTGTGATGGTGGTGTTCAATTCAGCATAATGGTGGTCTCATCCTCCCCGGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   665 GACTTGTGGGTGGTTGTGTTCCAGTTCAGCACATCATGTTGGCCTTATCCTGCCTGGT 724

Db   421 ATCGTCATCCTCCTGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   725 ATTGTCATCCTGTCCGTTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC 784

Db   481 CAGAAGCGCAAGGCCCTCAAGACAGCAGTCATCCTCATCCTCATCCTTTTGCCTGCTGG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   785 CAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTTTTCTTCGCCTGTTGG 844

Db   541 CTGCCATATTATGTGGGGATGGGGATCAGCATCGACTGACTCCTTTTGGGGAGTCATCAAGCAA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   845 CTGCCTTACTACATTGGGGATCAGCATCGACTCCTTCATCCTTCATCCTCCTCGAAATCATCAAGCAA 904

Db   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   905 GGGTGTGAGTTTGAGAACACGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 964

Db   661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
          ||||||||||||||||||||||||||||||

Qy   965 TTCCACTGTTGTCTGAACCCCATCCTCTAT 994

RESULT 15
US-10-151-274-14
; Sequence 14, Application US/10151274
; Publication No. US20030064071A1
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Tneo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10/151,274
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/517,605
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-151-274-14

Query Match   75.0%; Score 517.2; DB 13; Length 1679;
Best Local Similarity 85.8%; Pred. No. 1.5e-156;
Matches   592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
         |||||||||||||||||||||||||||||||||||||| || |||||||||||||||||

Db 320 CTGCACCTGTCAGTGGCCGACCTCCCTTGTCATCACACGCTTCCCTTCTGGGCAGTTGAT 379

Qy   61 GCCATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||

Db 380 GCCGTGGCAAACTGGTACTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA 439

Qy 121 GTCAACCTCCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCTGCC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 440 GTCAACCTTCTACAGCAGCGTCCTCATCCTGGCCTTCATCAGCCTGACCGGCTACCTGCC 499

```
Qy   181  ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
          | |||||||| ||||| |||||| |||||||||||||| ||||||||  ||  ||||||
Db   500  ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAGGCTGTTGGCTGAAAAGGTGGTCTAT  559

Qy   241  GTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
          ||  ||||||||||||  ||||||| || ||||  || |||||||||||||||||||| |
Db   560  GTTGGGGTCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCATCTTTGCCAACG--  617

Qy   301  AGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT  360
          |||||       ||||||| ||||||| || || |||  ||||||||||| ||||||
Db   618  ---------TCAGTGAGGCAGATGACAGCAGATATCTGTGACCGCCTTCTACCCCAAT  664

Qy   361  AGCCTGTGGATGGTGTGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC  420
          || |||| ||||| ||||| || ||||| |||| ||||| |||||| || ||||||||
Db   665  GACTTGTGGGTGGTTGTGTTCCAGTTTCAGCACACATCATGGTTGGCCTTATCCTGCCTGGT  724

Qy   421  ATCGTCATCCTCTCCTGTTACTGACATCATCATCCTAAGCTGTCACACTCCAAGGGCCAC  480
          ||||||||| || |||| ||| |||| ||||||| |||||||||||||||||||||||
Db   725  ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC  784

Qy   481  CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCGACATCCTTCATCCTTTGGGAGTCATCAAGCAA  540
          |||||||||||||||||||| || || ||||| |||||||| || | ||||| |||||| |
Db   785  CAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTGGCTTCTTCGCCCTGTTGG  844

Qy   541  CTGCCATATTATGTGGGGATCAGCATGACTCCTTCATCCTTTGGGAGTCATCAAGCAA  600
          |||||  | |  ||||| |||||||| || |||||| |||||||||| ||||||||||
Db   845  CTGCCTTACTACATTGGGATCAGCATCAGCATCCTCCTGGAAATCATCAAGCAA  904

Qy   601  GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGGCCCTCGCCTTC  660
          ||| ||| |||||||| | |||| |||||||| | ||| |||||||| |||||| |||||
Db   905  GGGGTGCGAGTTTGAGAACACTGCGACAAGTGCACAAGTGGATTTCCATCACCGAGGGCCCTAGCTTTC  964

Qy   661  TTCCACTGTGTTGCCTGAACCCCATCCTCTAT  690
          ||||||||||| ||||| ||||||||||||||
Db   965  TTCCACTGTTGTCTGAACCCCATCCTCTAT  994
```

Search completed: May 17, 2004, 22:14:04
Job time : 341.746 secs

```
          GenCore version 5.1.6
      Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      May 17, 2004, 05:50:31 ; Search time 1991.6 Seconds
                                       (without alignments)
                                       10345.919 Million cell updates/sec

Title:          US-09-367-052-3
Perfect score:  690
Sequence:       1 ctgcacctgtcagtggctga.........gcctgaacccatcctctat 690

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27513289 seqs, 1491090276 residues

Total number of hits satisfying chosen parameters:      55026578

Minimum DB seq length:   0
Maximum DB seq length:   2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
             1:  em_estba:*
             2:  em_esthum:*
             3:  em_estin:*
             4:  em_estmu:*
             5:  em_estov:*
             6:  em_estpl:*
             7:  em_estro:*
             8:  em_htc:*
             9:  gb_est1:*
            10:  gb_est2:*
            11:  gb_htc:*
            12:  gb_est3:*
            13:  gb_est4:*
            14:  gb_est5:*
            15:  em_estfun:*
            16:  em_estom:*
            17:  em_gss_hum:*
            18:  em_gss_inv:*
            19:  em_gss_pln:*
            20:  em_gss_vrt:*
            21:  em_gss_fun:*
            22:  em_gss_mam:*
            23:  em_gss_mus:*
            24:  em_gss_pro:*
            25:  em_gss_rod:*
            26:  em_gss_phg:*
            27:  em_gss_vrl:*
            28:  gb_gss1:*
            29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result      %     Query
No.   Score Match Length DB ID                     Description
----- ----- ----- ------ -- --------  ----------------------------------
  1   517.2  75.0   872  14 CD251331  CD251331 AGENCOURT
  2   511.6  74.1  1201  13 BX421287  BX421287 BX421287
  3   506.2  73.4  1201  13 BX358346  BX358346 BX358346
  4   505.4  73.2  1201  13 BX421511  BX421511 BX421511
```

```
  5  494.8 71.7 1201 13 BX396365  BX396365 BX396365
  6  492.2 71.3 1091 13 BX397685  BX397685 BX397685
  7  491.8 71.3 1201 13 BX353253  BX353253 BX353253
  8  484   70.1 1051 12 BM545259  BM545259 AGENCOURT
  9  476.4 69.0 1201 13 BX462627  BX462627 BX462627
 10  468   67.8  906 13 BU523025  BU523025 AGENCOURT
 11  453.8 65.8 1201 13 BX418530  BX418530 BX418530
 12  436.8 63.3 1201 13 BX462182  BX462182 BX462182
 13  435.6 63.1  749 14 CD466324  CD466324 LeukoN2_3
 14  435   63.0  582 12 BG145042  BG145042 ut75a01.y
 15  431.8 62.6  952 13 BX362477  BX362477 BX362477
 16  430.6 62.4  817 14 CD520635  CD520635 AGENCOURT
 17  419.6 60.8  782 14 CB985366  CB985366 AGENCOURT
 18  419   60.7  846 10 BF338608  BF338608 602034286
 19  418   60.6  446  9 AI131863  AI131863 uc36h08.r
 20  411.2 59.6  945 13 BX443258  BX443258 BX443258
 21  409.4 59.3  750 12 BG613352  BG613352 602641234
 22  409   59.3  911 13 BX443334  BX443334 BX443334
 23  404.6 58.6  910 12 BI762229  BI762229 603049139
 24  397.8 57.7  891 12 BI824663  BI824663 603033815
 25  397.2 57.6  793  9 AA182270  AA182270 mt83f01.r
 26  393.2 57.0  886 13 BX443228  BX443228 BX443228
 27  387.4 56.1  950 13 BQ718617  BQ718617 AGENCOURT
 28  386   55.9  612 12 BG483197  BG483197 602503140
 29  376.2 54.6  528 12 BG277031  BG277031 uv10e03.y
 30  376.2 54.5  553 12 BG145062  BG145062 ut75b11.y
 31  375.6 54.4  928 14 CD389217  CD389217 AGENCOURT
 32  370   53.6  827 12 BI761664  BI761664 603046395
 33  368.4 53.4 1034 12 BM920800  BM920800 AGENCOURT
 34  366   53.0  600 13 BU919690  BU919690 6033-58 M
 35  362.2 52.5  765 12 BM387369  BM387369 UI-R-CN1-
 36  361.8 52.4  876 14 CA454255  CA454255 AGENCOURT
 37  359   52.0  639  9 AA816049  AA816049 vr14h02.r
 38  354.8 51.4  532 10 AW388109  AW388109 MR4-ST012
 39  339   49.1  875 12 BI756157  BI756157 603030061
 40  334.2 48.4  499 10 AW389565  AW389565 RC2-ST016
 41  333.2 48.3  581 14 CF540844  CF540844 UI-M-GW0-
 42  331.2 48.0  976 13 BU146456  BU146456 AGENCOURT
 43  328.4 47.6  958 12 BG174412  BG174412 602334232
 44  324.8 47.1  444 12 BP430205  BP430205 BP430205
 45  323.8 46.9  657 13 BU283174  BU283174 603863101
```

ALIGNMENTS

```
RESULT 1
CD251331
LOCUS       CD251331      AGENCOURT 14211919 NIH MGC 179 Homo sapiens cDNA clone
DEFINITION  AGENCOURT 14211919 NIH_MGC_179 Homo sapiens cDNA clone
            IMAGE:30384274 5', mRNA sequence.
ACCESSION   CD251331
VERSION     CD251331.1  GI:31011797
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 872)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
Plate: NDAM449  row: h  column: 11
High quality sequence stop: 654.
```

FEATURES
source   1..872
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:30384274"
              /tissue_type="Pituitary"
              /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
              /clone_lib="NIH_MGC_179"
              /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
              (destroyed); Site_2: NotI; Library is oligo-dT primed and
              directionally cloned (EcoRV site is destroyed upon
              cloning). Average insert size 1.1 kb. Library was
              constructed by (Invitrogen). Note: this is a NIH_MGC
              Library."

ORIGIN

Query Match   75.0%; Score 517.2; DB 14; Length 872;
Best Local Similarity   85.8%; Pred. No. 2.7e-111;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
        |||||||||||||||| ||||||||||||||||||||| || ||||||||||||||||||
Db   59 CTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT 118

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACT 120
        ||| ||| ||||||||||||||||| || | ||| |||| |||| ||||| || ||
Db 119 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCATGTCATCTCACA   178

Qy 121 GTCAACCTCTACAGCACCAGCGGTTCCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
        |||||||||||||| ||||| | |||| ||||||||||||||||||||||||||||||| |||
Db 179 GTCAACCTCTACAGCAGCAGTGTCCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCC 238

Qy 181 ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
        || |||||||||||||||||||| | ||||||||| |||||||||||| || ||||||
Db 239 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGTGTTGGCTGAAAAAGGTGGTCTAT 298

Qy 241 GTGGGGCGTCTGGATCCCCAGCCCTCCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
        || |||||||||||||||| ||||||||||||||||||||||||||||||||||| |
Db 299 GTTGGCGTCTGGATCCCCTGCCCTCCCTGACTATACCTGACTTCATCTTTGCCAACG-- 356

Qy 301 AGCCAGGGGGACACTCAGTCAGGGGATGACAGGTACATCTGTGACCGGCCTTTACCCCGAT 360
                        ||||| |||||||||||||| |||||||||||||||||| |||||| |
Db 357 --------TCAGTGAGGCAGATGACAGATATATCTGTGACCGGCCTTCTACCCCAAT   403

Qy 361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGGTCTCATCCTGGCCCCGGC 420
        | | || ||| ||||| ||||| |||||| || ||||| ||||||||||||
Db 404 GACTTGTTGGGTGGTGTTGTTCAGTTTCAGCACCATCATGGGTGGCCTTATCCTGGCCTGGT 463

Qy 421 ATCGTCATCTCCTGTACTGCATCATCATCCTAAGTGTCACACTCCAAGGGGCCAC   480
        || ||||| | | | ||||| |||||||| | |||||||||||||||||||
Db 464 ATTGTCATCCGTCCGCTATTGCATTATCATCCTCCAAGCTGTCACACTCCAAGGGGCCAC   523

Qy 481 CAGAAGGCCAAGGCCCTCAAGACAGCAGTCATCCTCATCCTCTTCTTTGCCTGCTGG   540
        |||||||||||||||||||||||||| | |||||||||||||| |||| |||| |||
Db 524 CAGAAGGCCAAGGCCCTCAAGACAGCCAGCAGTCATCCTCATCCTGGCTTTCTTCGCCTGTTGG 583

Qy 541 CTGCCATTATGTGGGGATCAGGATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA 600
        ||||| | |||||||||| ||||||||||||||||||||||| |||||| ||||||||||
Db 584 CTGCCTTACACATTGGGATCAGCATCGACTCCTTCATCCTTCATCCTGGGAAATCATCAAGCAA 643

Qy 601 GGATGTGACTTCGAGAGCATTGTGCACAAGTTGGATTTCCATCACCGAGGCCCTCGCCTTC 660
        || ||| | || |||| || |||||||||| |||||||||||||||||||||| || |||
Db 644 GGGTTGGCGTCTGGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTTC 703

Qy 661 TTCCACTGTTGCCTGAACCCCATCCTCTAT 690
        ||||||||||| |||||||||||||| |||
Db 704 TTCCACTGTTGTCTGAACCCCATCCTTTAT 733
```

RESULT 2
BX421287

LOCUS   BX421287   1201 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION BX421287 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG004YG09 5-PRIME, mRNA sequence.
ACCESSION   BX421287
VERSION   BX421287.1 GI:30766068
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DG004AD05QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DG004AD05QP1.

FEATURES   Location/Qualifiers
     source   1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG004YG09"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

ORIGIN

Query Match   74.1%; Score 511.6; DB 13; Length 1201;
Best Local Similarity   84.6%; Pred. No. 6.6e-110;
Matches 584; Conservative 6; Mismatches 85; Indels 15; Gaps 1;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT   60
        |||||||||||||||| |||||||||||||||||||| || ||||||||||||||||||
Db  334 CTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT  393

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACT 120
        |||  ||| ||||||||||||||||| ||  | ||| |||| |||| ||||| || ||
Db  394 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCATGTCATCTCACA   453

Qy  121 GTCAACCTCTACAGCACCAGCGGTTCCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
        |||||||||||||| ||||| | |||| ||||||||||||||||||||||||||||||| |||
Db  454 GTCAACCTCTACAGCAGCAGTGTCCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCC 513

Qy  181 ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT  240
        || |||||||||||||||||||| | |||||||||  |||||||||||| || ||||||
Db  514 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGTGTTGGCTGAAAAAGGTGGTCTAT 573

Qy  241 GTGGGGCGTCTGGATCCCCAGCCCTCCCTGACTATACCTGACTTCATCTTTGCCGACGTC  300
        || |||||||||||||||| |||||||||||||||||||||||||||||||||||  |
Db  574 GTTGGCGTCTGGATCCCCTGCCCTCCCTGACTATACCTGACTTCATCTTTGCCAACG--  631

Qy  301 AGCCAGGGGGACACTCAGTCAGGGGATGACAGGTACATCTGTGACCGGCCTTTACCCCGAT 360
                        ||||| |||||||||||||| |||||||||||||||||| |||||| |
Db  632 --------TCAGTGAGGCAGATGACAGATATATCTGTGACCGGCCTTCTACCCCAAT    678

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGGTGGGTCCATCCTGGCCCCGGC 420
```

RESULT 3

```
Db   679 GACTTGTGGGTGGTTGTTGTTCCAGTTGTGTTCAGCACATCATGGTTGGCTTATCCTGCTGGT  738
          ||||| |||| |||  |||   ||  || |||||  ||| |||| ||  ||| |||||||||||
Qy   421 ATCGTCATCCTCCTCCTGTTACTGCATCATCATCTTCTTAAGCTGTCACACTCCAAGGGCCAC  480
          ||||| ||||||| |||||| |||||||||||||||||||||||||||||||||||||||
Db   739 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACTCCAAGGGCCAC  798
Qy   481 CAGAAGCGCAAGGCCCTCAAGACGACACCCTAGTCATCCTCATCCTTCTTGCCTGCTGG  540
Db   799 CAGAAGCGCAAGGCCCTCAAGACCCAGTCATCCTCATCCTTCTTCGCCTGCGTGTGG  858
Qy   541 CTGCCATATTGTGGGGATCAGCATCGATCCTCCTTCATCCTTTCTTTGCCTGCTGG  600
Db   859 CTGCCTTACTACATTGGATCAGCATCGATCCTGACCTCCTTCATCCTTTCATCAAGCAA  918
Qy   601 GGATGTGACTTCGAGAGCATTGGGATCCATCCTCCATCCTGCGCCCTTC  660
Db   919 GGGTGTGAGTTTGAGAACACTGGGCACCGGATTYCATCACCGAGGCCCTAGTTTY  978
Qy   661 TTCCACTGTTGCCTGAACCCATCCTCTAT  690
Db   979 TTYCACTGTGTYTGAACCCCWTCYTCTAT 1008
```

RESULT 3
BX358346

ORIGIN

Query Match   73.4%;   Score 506.2; DB 13; Length 1201;
Best Local Similarity 85.3%;   Pred. No. 1.2e-108;
Matches 580;   Conservative   1;   Mismatches   85;   Indels   14;   Gaps   1;

```
Qy     1 CTGCACCTGTCAGTGGCTGACCTCCCTTTTGTCATCACACTCCCCTTCTGGGCAGTTGAT    60
         ||||||||||||||||| ||||||||| ||||||||| || ||||||||||||||||||
Db   364 CTGCACCTGTCAGTGGGCCGACCTCCCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT   423
Qy    61 GCCATGGCTGACGGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACT  120
         |||  |||| ||  |||||||||||| ||  | ||| ||||| |||||| |||||||||
Db   424 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA  483
```

RESULT 4
BX421511

```
Qy   121 GTCAACCTCTACAGCAGCGGCGTTCTCATCCTGGCCTTCATCAGCGGTGGACCGGTACCTCGCC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 GTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCAGTCAGTGGTGGACCGGCTACCTGGCC  543
Qy   181 ATTGTCCAGGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGGAAAAGGCAGTCTAT  240
Db   544 ATTGTCCAGGCCACCAACAGTCAGAGGCCAAGGAAGCTGTGTGGTGGCTGAAAAGGTGTCTAT  603
Qy   241 GTGGGCGTCGGATCCCAGGCCTCCTCCTGACTATACCGACTTCATCTTTGCGCAGTC  300
Db   604 GTTGGCGTTGGGATCCCTGCCCTCCTGACTATTCCCGACTCCATCTTTGCCAAGTC  663
Qy   301 AGCCAGGGGGACATCAGTCAGGGGGATGAGCAGGTACATCGTGACGTACATCGTGACCCCGAT  360
Db   664 AGTGA----------NGGCAGATGACAGATATATCGTGTGACCGCGCTTCTACCCCAAT  709
Qy   361 AGCCTGTGGATGGGTGGTGTTCAATTCAGCATTAAATGGTGGGTCTCATCCTGCCCGGC  420
Db   710 GACTTGTGGGTGGTGTGTGTGCCAGTTCAGCACATCAGTCATCATCTCCAAGCTGCTGGT  769
Qy   421 ATCGTCATCCTCCTCCTGTTACTGCATCATCATCTTCTTAAGCTGTCACACTCCAAGGGCCAC  480
Db   770 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACTCCAAGGGCCAC  829
Qy   481 CAGAAGCGCAAGGCCCTCAAGACGACAGCAGTCATCCTCATCCTTCTTGCCTGCTGG  540
Db   830 CAGAAGCGCAAGGCCCTCAAGACCCAGTCATCCTCATCCTTCTTCGCCTGCGTGTGG  889
Qy   541 CTGCCATATTATGTGGGGATCAGCATCAGCATCAGCTCCTTTCATCCTTTGGGAGTCATCAAGCAA  600
Db   890 CTGCCTTACTACATTGGATCAGCATCAGCATCGGATCTCCATCCTCCTGGAAATCATCAACCAA  949
Qy   601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGAGGCCCTCGCCTTC  660
Db   950 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 1009
Qy   661 TTCCACTGTTGCCTGAACCC  680
Db  1010 TTCCACTGTGTYTGAACCC  1029
```

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 9106 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI015DF10QP1&cluster=i1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI015DF10QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI015YL20"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

```
Query Match             71.7%;  Score 494.8;  DB 13;  Length 1201;
Best Local Similarity   85.3%;  Pred. No. 6.1e-106;
Matches  580;  Conservative   1;  Mismatches  83;  Indels  16;  Gaps   2;

Qy      1 CTGCACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT    60
          ||||||||||||||||| |||||||||||||||||| || |||||||||||||||||||
Db    350 CTGCACCTGTCAGTGGCCGACCTCCTTTGTCATCACGCTTCCCTTCTCGGGCAGTTGAT   409

Qy     61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
          |||  || |||||||||||||||||||   | | || ||||| |||||| ||||||||||
Db    410 GCCGTGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA   469

Qy    121 GTCAACCTTCTACAGCAGCGGTTCCATCTCGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
          |||||||||||||||||||| ||||| |||||||||||||||| |||||||||||||| ||
Db    470 GTCAACCTTCTACAGCAGCGTCCATCCTCGGCCTTCATCAGTCTGGACCGGTACCTCGGCC  529

Qy    181 ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
          || ||||||||||||| |||||| ||||||||||| ||||| |||||||||| ||||||
Db    530 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGGTGGCTGAAAAGGTGGTCTAT   589

Qy    241 GTGGGGCTCGATCCCAGCCCTCCTCCTGCTACTATTCCGACTTCATCTTTGCCGACGTC   300
          || || ||||||||| ||||||||||||||| ||||||||| ||||||||||||||| ||
Db    590 GTTGGGCTCGATCCCGATCCCGCTGCTACTATTCCGACTTCATCTTTGCCAACG--     647

Qy    301 AGCCAGGGGGACATCAGTCAGGGGATGACAGAGTACATCTGAGATATATCTGTGACCGAT   360
                   ||||| |||||| ||||||||||| ||| |||||||||||||||| |||||||
Db    648 ----------TCAGTAGGGCAGATGACAGATATATCTGTGACCGCCTTCTACCCCAAT    694

Qy    361 AGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGGTCTCATCCTGCCCGGC   420
          ||||| ||| ||||| || |||||||||||| ||| |||| |||| |||||||||| ||
Db    695 GACTTGTGGGTGGTTGGTGTTCCAGTTTCAGCACACATCATGGTTGGCCTTATCCTGCTTGGT   754

Qy    421 ATCGTCATCCTCTCCGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCAC   480
          || ||||||||| ||| | ||||||||||||| |||||||||||||||||||||||||
Db    755 ATTGTCATCCTGTCCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC   814

Qy    481 CAGAAGGCAAGGCCCTCAAGACGACGAGTCATCCATCCTAGCTTTCTTGCCTGCTGG   540
          |||||||||||||||||||||||| || ||||||||||| |||||||||||||||||
```

```
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG007YM12"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6;1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN
```

```
Query Match             73.2%;  Score 505.4;  DB 13;  Length 1201;
Best Local Similarity   85.0%;  Pred. No. 1.9e-108;
Matches  578;  Conservative   3;  Mismatches  85;  Indels  14;  Gaps   1;

Qy      1 CTGCACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT    60
          ||||||||||||||||| |||||||||||||||||| || |||||||||||||||||||
Db    358 CTGCACCTGTCAGTGGCCGACCTCCTTTGTCATCAGCGCTTCCCTTCTCGGGCAGTTGAT   417

Qy     61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
          |||  || |||||||||||||||||||   | | || ||||| |||||| |||||||||
Db    418 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCMAGGCAGTCCATGTCATCTACACA   477

Qy    121 GTCAACCTTCTACAGCAGCGGTTCCATCTCGGCCTTCATCAGCCTGGACCGGTACCTCGCC  180
          |||||||||||||||| |  ||||| ||||||||||||||||| |||||||| |||||| ||
Db    478 GTCAACCTTCTACAGCMGTGTCCATCCTCGGCCTTCATCAGTCTGGACCGCCTACCTGGCC   537

Qy    181 ATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT   240
          || ||||||||||||| |||||| ||||||||||| ||||| |||||||||| ||||||
Db    538 ATCGTCCACGCCACCMACAGTCAGAGGCCAAGGAAGCTGGTGGCTGAAAAGGTGGTCTAT   597

Qy    241 GTGGGCGTCTGGATCCCAGCCCCTCCTCCTGCTGCTATACCTGACTTCATCTTTGCCGACGTC   300
          || || |||||||||| ||||| ||||||||||| ||||| |||||||||||||| |||| ||
Db    598 GTTGGCGTCTGGATCCCCTGCTCCTGCTGCTGCTATTCCCGACTTCATCTTTGCCAACGTC   657

Qy    301 AGCCAGGGGGGA------------------                               360
          |||  | || |
Db    658 AGTGA-----NGGGCAGATGACAGATATATCTGTGACCGTGACCGCCTTCTACCCCAAT   703

Qy    361 AGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGGTCTCATCCTGCCCGGC   420
          ||||| ||| ||||| || |||||||||||| ||| |||| |||| ||||||||| ||
Db    704 GACTTGTGGGTGGTGGTGTTCCAGTTTCAGCACACATCATGGTTGGCCTTATCCTGCCTGGT   763

Qy    421 ATCGTCATCCTCTCCGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCAC   480
          || ||||||||| ||| | ||||||||||||| |||||||||||||||||||||||||
Db    764 ATTGTCATCCTGTCCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC   823

Qy    481 CAGAAGGCAAGGCCCTCAAGACGACGAGTCATCCATCCTAGCTTTCTTGCCCTCGCCTTC   540
          ||||||| ||||||||||||||| || ||||||||||| |||||||||| ||||||||||
Db    824 CAGAAGGCCAAGGCCCTCAAGACCACCAGTCATCCATCCTCATCCTGCTTCTTCGCCTTC   883

Qy    541 CTGCCATATTATGTGGGGATCAGCATCGACTTCCTTCCATCCTTTGGGAGTCATCAAGCAA   600
          ||||| || || ||||||||||||||||| || |||||||| || ||||| |||||||||
Db    884 CTGCCTTACTACATTGGGATCAGCATCGAGCTCCTTCCATCCCTGGAAATCATCAAGCAA   943

Qy    601 GGATGTGACTCGAGAGACATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCCTC   660
          || |||||| | ||| ||| | |||||||||||||| |||||||| |||||||| || ||
Db    944 GGGGTGTGAGTTTGAGAAGACACTGTGCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC  1003

Qy    661 TTCCACTGTTGCCTGAACCC   680
          ||||||||||| ||||||||
Db   1004 TTCCACTGTTGTCTGAACCC  1023
```

```
RESULT 5
BX396365
LOCUS       BX396365        1201 bp    mRNA    linear   EST 13-MAY-2003
DEFINITION  BX396365 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI015YL20 5-PRIME, mRNA sequence.
ACCESSION   BX396365
VERSION     BX396365.1  GI:30609085
```

```
Qy  181 ATTGTCCACGCCACCACCAGTCATCCTCATCCTTCTTCGGCCAGTCTAT 240
            |||||||||||||||||||||||||| |||||||||||||||||||||
Db  532 ATCGTCCACGCCACCACCAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCGTAT 591

Qy  241 GTGGGCGTCTGGATCCCAGCCCTCCTGCCTCCTGCTATACCTGACTTCATCTTTGCGACGTC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  592 GTTGGGCGTCTGGATCCTGCCCTCCTGCTGTGACTTCCGACTATTCCGACTTCATCTTTGCCAACGTC 651

Qy  301 AGCCAGGGGACATCAGTCAGGGGATGACAGTGACATCTGTGACCGCCTTTACCCCGAT 360
                       ||||||||||||||||||||||
Db  652 ---NGGCAGATGACAGATATATCGTGACATCTGTGACCGCCTTCTACCCCAAT 697

Qy  361 AGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGTCTCATCCTGCCGGC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  698 GACTTGTGGGGTGGTTGTTGTTCCAGTTTCAGCACACATCATGGTTGGCCTTATCCTGCCTGGT 757

Qy  421 ATCGTCATCCTCTCCGTTACTGCACTCATCATCATCTCTTAAGCTGTCACACTCCAAGGGGCAC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  758 ATTGTCATCCTGTCCTCCTATGCATTATCATCTCCAAGTCGTCACACTCCAAGGGGCAC 817

Qy  481 CAGAAGCGCAAGGCCCTCAAGACGACAGACAGTCATCATCCTCATCCT-AGCTTTCTTTGCCTGCT- 538
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  818 CAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCYTCATCGTCGGGGCTTTCTTCGCCTGTTK 877

Qy  539 GGCTGCCATATTATGTGGGGATCAGCATGCAGATCGGACTCCTTCATCCTTTTGGGAGTCATCAAGC 598
            ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||
Db  878 GGCTGCCTTACTACATTGGGATCAGGCATCAGCATCCTTCATCCTTCATCCTCCTGGAAATCATCAAGC 937

Qy  599 AAGGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCT 658
            ||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||
Db  938 AAGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGGATTTCCATCACCGAGGCCCTAGCTT 997

Qy  659 TCTTCCACTGTGCTGCGAACCCCATCCTCTAT 690
            |:|||||||||||||||||||||||||||||
Db  998 TYTTCCACTGTTGTTTGAACCCCATCCTCTAT 1029
```

RESULT 7

BX353253
LOCUS   BX353253   Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED   1201 bp   mRNA   linear   EST 02-MAY-2003
DEFINITION BX353253 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC003YG03 5-PRIME, mRNA sequence.
ACCESSION   BX353253
VERSION   BX353253.1   GI:30347726
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DC003AD02QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DC003AD02QP1.
FEATURES   Location/Qualifiers
source   1..1201
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DC003YG03"
            /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

```
Db  815 CAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTGGCTTTCTTCGGCTGTTGG 874

Qy  541 CTGCCCATATTATGTGGGGGATCAGCTCCTTCATCCTTTTTGGGAGTCATCAAGCAA 600
            || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  875 CTGCCTTACTACATTGGGATCAGCATCAGCATCCTTCATCCTTCATCCTCTGGGAAATCATCAAGCAA 934

Qy  601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
            || ||||||||||||||| |||||||||||||||| ||||||||| ||||||||||||
Db  935 GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGGAT-TYCATCACCGAGGCCCTAGCTTTC 993

Qy  661 TTCCACTGTGCTGCGAACCC 680
            |||||||||| ||||||
Db  994 TTCCACTGTGTGTCTGAACCC 1013
```

RESULT 6

BX397685
LOCUS   BX397685   Homo sapiens PLACENTA COT 25-NORMALIZED   1091 bp   mRNA   linear   EST 13-MAY-2003
DEFINITION BX397685 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI041YO13 5-PRIME, mRNA sequence.
ACCESSION   BX397685
VERSION   BX397685.1   GI:30621359
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1091)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI041AH07QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI041AH07QP1.
FEATURES   Location/Qualifiers
source   1..1091
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DI041YO13"
            /tissue_type="PLACENTA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and cloned into the Not I and EcoR V
            sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match   71.3%;   Score 492.2; DB 13; Length 1091;
Best Local Similarity 85.0%;   Pred. No. 2.4e-105;
Matches 588;   Conservative 2;   Mismatches 86;   Indels 16;   Gaps 3;
```

```
Qy    1 CTGCACCTGCACGGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTCGGGCAGTTGAT 60
            |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  352 CTGCACCTGTCAGTGGCCGACCCTCCTCTTTGTCATCACAGCTTCCCCTTCCTCTGGGCAGTTGAT 411

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTCACACT 120
            |||||||||||||| ||||||||||||||||||| ||||||||||||||||||||||
Db  412 GCCGGGCAAACTGTACTTTGGGAACTTCCTATGCAAGGCAGCTCCATGTCATCATCTACACACA 471

Qy  121 GTCAACCTTCTACAGCAGCGTTCTTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCC 180
            ||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||
Db  472 GTCAACCTTCTACAGCAGTGTCCTTCATCCTGGCCTTCATCAGTCTGACCGGTACCTCGCC 531
```

TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
         Email: cgaps-r@mail.nih.gov
         Tissue Procurement: Invitrogen
         cDNA Library Preparation: Life Technologies, Inc.
         cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
         DNA Sequencing by: Agencourt Bioscience Corporation
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLAM12720 row: e column: 12
         High quality sequence start: 20
         High quality sequence stop: 753.

FEATURES
     source
         Location/Qualifiers
         1..1051
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:5726963"
         /tissue_type="hippocampus"
         /lab_host="DH10B"
         /clone_lib="NIH_MGC_124"
         /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV
         (destroyed); Site_2: NotI; RNA source male hippocampus,
         age 27. Library is oligo-dT primed and directionally
         cloned (EcoRV site is destroyed upon cloning). Average
         insert size 1.4 kb, insert size range 0.9-4 kb. Library is
         normalized and enriched for full-length clones and was
         constructed by C. Gruber (Invitrogen). Research Genetics
         tracking code 012."

ORIGIN

Query Match 70.1%;   Score 484;   DB 12;   Length 1051;
Best Local Similarity 84.5%;   Pred. No. 2e-103;
Matches 585;   Conservative 0;   Mismatches 90;   Indels 17;   Gaps 3;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCCTTCTGTCATCACACCCTCTTGTCATCACACT CTGGGCAGTTGAT  60
          ||||||||||||||||| ||||||   |||||||||| ||  ||||||||         |||||||||||
Db  299 CTGCACCTGTCAGTGGCCGACCTCTCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT            358

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT           120
          |||.||||   |||||||||||||||  ||  ||  ||||||||  |||||||| || ||
Db  359 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGCAGTCATGTCATCCACACA            418

Qy  121 GTCAACCTTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC            180
          ||||||||.||||||| |||| ||||||||||||||||| |||||||||||||||| ||
Db  419 GTCAACCTTCACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGTACCTCGGCC        478

Qy  181 ATTGTCCAGCCACCAAACTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT           240
          ||.||||||||||||| |   .|||||||||||.|||||||| |||||.||| |||||
Db  479 ATCGTCCAGCCACCAACAGTCAGAGGCCAAGGAAGCTGGCTGGTGAAAAAGGTGTCTAT         538

Qy  241 GTGGGCGTCTGGATCCCAGCCCTCCTGCTGTCTGACTTATCCTCATCTTTGCCGACGTC          300
          ||.||||||||||||||   ||||||||||||||||| ||||| ||||||||||.||||
Db  539 GTTGGGCGTCTGGATCCCGTCCTGCTGTCTGACTATTCCCGACTTCATCTTTGCCAACG--        596

Qy  301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGCAGATACAGGTACATCGTGACCGCCTTACCCCGAT   360
                 |||||||||||  |||||||| ||| ||||||||||||||||||||||||
Db  597 -------TCAGTGAGGCAGATGACAGATATCGTGACCGCCTTCTACCCCCAAT                643

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGC           420
          |||||||||| ||||||||||||||||||||  |||| ||  ||||  ||||||||
Db  644 GACTTGTGGGGTGGTTGTGTTTCCCAGTTCAGCACACTCATGGTTGCCTTATCCTGCCTGGT      703

Qy  421 ATCGTCATCCTCCTGTTACTGCATCATCATCCTTAAGCTGTCACACTCCAAGGGGCCAC          480
          ||.|||||||| ||| ||||||||.|||||||||.|||||||||||||||||||||||||
Db  704 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGGCCAC        763

Qy  481 CAGAGAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTCTTGCCTGCTGTGG        540
          |||||||||||||||||||||||| |||||||||||||||||| || ||| |||| || ||
Db  764 CAGAGAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTGGTTTTCTTCGCCTGTTGG       823
```

ORIGIN

Query Match 71.3%;   Score 491.8;   DB 13;   Length 1201;
Best Local Similarity 84.9%;   Pred. No. 3.1e-105;
Matches 577;   Conservative 2;   Mismatches 85;   Indels 16;   Gaps 2;

```
Qy    1 CTGCACCTGTCAGTGGCTGACCTCCCTTCTGTCATCACACCCTCTTGTCATCACACT CTGGGCAGTTGAT  60
          ||||||||||||||||| ||||||   |||||||||| ||  ||||||||         |||||||||||
Db  367 CTGCACCTGTCAGTGGCCGACCTCTCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT            426

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT           120
          |||.||||   |||||||||||||||  ||  ||  ||||||||  |||||||| || ||
Db  427 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGCAGTCATGTCATCCACACA            486

Qy  121 GTCAACCTTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC            180
          ||||||||.||||||| |||| ||||||||||||||||| |||||||||||||||| ||
Db  487 GTCAACCTTCACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGTACCTCGGCC        546

Qy  181 ATTGTCCAGCCACCAAACTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT           240
          ||.||||||||||||| |   .|||||||||||.|||||||| |||||.||| |||||
Db  547 ATCGTCCAGCCACCAACAGTTAGAGGCCAAGGAAGCTGGCTGGTGAAAAAGGTGTCTAT         606

Qy  241 GTGGGCGTCTGGATCCCAGCCCTCCTGCTGTCTGACTTATCCTCATCTTTGCCGACGTC          300
          ||.||||||||||||||   ||||||||||||||||| ||||| ||||||||||.||||
Db  607 GTTGGGCGTCTGGATCCCGTCCTGCTGTCTGACTATTCCCGACTTCATCTTTGCCAACG--        664

Qy  301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCGTGACCGCCTTCTACCCCGAT         360
                 ||||| |||||  |||||||| ||| |||||||||||||||||||||||
Db  665 -------TCAGTANGGCAGATGACAGATATCTGTGACCGCCTTCTACCCCCAAT               711

Qy  361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGC          420
          |||||||||| ||||||||||||||||||||  |||| ||  ||||  ||||||||
Db  712 GACTTGTGGGGTGGTTGTGTTTCCCAGTTCAGCACACTCATGGTTGCCTTATCCTGCCTGGT      771

Qy  421 ATCGTCATCCTCCTGTTACTGCATCATCATCCTTAAGCTGTCACACTCCAAGGGCCAC           480
          ||.|||||||| ||| ||||||||.|||||||||.||||||||||||||||||||||||
Db  772 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAC        831

Qy  481 CAGAGGCCAAGGCCCTCAAGACGACAGACAGTCATCCTCATCCTAGCTTTCTTGCCTGCTGG       540
          |||||  ||||||||||||||| ||||   ||||||||||||||| ||| |||.||| |||
Db  832 CAGAGAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTGGTTTCTTCGCCCTGTCGG       891

Qy  541 CTGCCATATTATGTGGGGATCCGACTCCCTTTGGGAGTGCATCAAGCAA                   600
          ||||| ||.|| |||||||||.|| ||| | |||||||.|||||||||
Db  892 CTGCCTTACTACATTGGGATCAGCAGCATCGACTCCTTCATCCTTCCTGAAAATCATCAAGCAA    951

Qy  601 GGATGTGACTTCGAGAGCAATTGTCACAGTGGATCCCATCACAGAGGGCCCTCGCCTTC          660
          || |||||.|| ||||| |  ||.|||||||||| .||||||.|| |||||||| |||.|
Db  952 GGGTGTGAGTTTGAGACACTGTGCACAGTGGATTTCCATCACCGCCAGCCCTAGCTTTTC       1011

Qy  661 TTCCACTGTGTTGCTGACCC  680
          || |||||||||  ||||||
Db 1012 TT-CACTGTGTTCTGAACCC  1030
```

RESULT 8
BM545259

Qy 541 CTGCCATATATGTGGGGATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCATCAAGCAA 600
         |||| |||||||||||||| |||||| ||||| ||||||||||||||||||||||||||
Db 824 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTTCCTGGAAATCATCAAGCAA 883

Qy 601 -GGATGGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGA-GGCCCCTCGCCT 658
          |||| ||| || |||| || || ||||||||||||||||||| | ||||||| |||||
Db 884 GGGGGTGTGAGTTTGAGAACACGTGCACAAGTGGATTTCATCACCGAGGGCCCTAGCTT 943

Qy 659 TCTTCACTGTTGCCTGAACCCCATCCTCTAT 690
         ||||||||||||||||| ||| |||||||||
Db 944 TCTTCACTGTTTGGCTGAACCCCATCCTCTAT 975

RESULT 9
BX462627
LOCUS   BX462627   1201 bp   mRNA   linear   EST 22-MAY-2003
DEFINITION BX462627 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
            clone CS0DH005YK10 5-PRIME, mRNA sequence.
ACCESSION   BX462627
VERSION   BX462627.1 GI:31029445
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DH005BF05QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DH005BF05QP1.
FEATURES   Location/Qualifiers
     source   1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DH005YK10"
                     /tissue_type="T CELLS (JURKAT CELL LINE)"
                     /cell_line="JURKAT CELL LINE"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

Query Match   69.0%; Score 476.4; DB 13; Length 1201;
Best Local Similarity   84.7%; Pred. No. 1.3e-101;
Matches 576;   Conservative   0;   Mismatches   86;   Indels   18;   Gaps   3;

Qy   1 CTGCACCTGGCTCAGTGGCTACTTTGGGGAAATTTTGTGTAGGCGTGTCCATATCATCACACT 60
         |||||||||| |||||||||||| || |||| |||||||| ||||||||||| || |||
Db 364 CTGCACCTGTCAGTGGCGCACCTTCCCTTTGTCATCACCGCTTCCCTTCTGGGCAGTTGAT 423

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGCCATATCATCTACACT 120
         ||||| |||| |||| ||||||||||||||||||| ||||||||||| ||| |||||
Db 424 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA 483

Qy 121 GTCAACCTCTACAGCAGCGTGTCCTCATCCTGGCCTTCATCACCTGGCCGTACCTCGCC 180
         ||||||||||||||| || ||| |||||||||| |||||||||| ||| |||||||
Db 484 GTCAACCTCTACAGCAGCGTGTCCTCATCCTGGCCCTTCATCAGTCTCAGTCTGGACCGGCC 543

Qy 181 ATTGTCCACGCCACCAAAGTCAAGGCCAAGGAAAACTGCTGGCTGAAAAGGCAGTCTAT 240
         |||||||||||||||| |||||||||||||| |||| ||||||||| |||| |||||
Db 544 ATTGTCCACGCCACCAAAGTCAAGGCCAAGGAAGTTGGCTGAAAAGGTGGTCTAT 603

Qy 241 GTGGGCGTCTGGATCCCAGCCCTCCTCCTGCCCTCCTCCTGGCTCTATACCTGACTTCATCGTC 300
         |||| |||||||||||||||||||||||| |||||| ||||||||| ||||||
Db 604 GTTGGCGTCTGGATCCCGTCCCTCCTGCCCTGACTTCCCGACTTCATCGTC 661

Qy 301 AGCAGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCGCCTTTACCCCGAT 360
         ||||||||||||| |||||||||||||||||||||||||||||||
Db 662 --------TCAGTGAGGCAGATGACAGATATCGTGACCGGCTTCTACCCCAAT 708

Qy 361 AGCCTGGATGGTGGTGTTCAATTCCAGCACATATAATGGTGGTCTCCATCCTGCCCGGC 420
         ||||||||||| |||||| ||||||||| ||| ||||||||||||||||||||||
Db 709 GACTTGTGGGTGGTGGTGTTCCAGTTTCAGCAACATATGGTGGCCTTATCCTGTCGGT 768

Qy 421 ATCGTCATCCTCTCCTGTTACTGCATCATCATCCTCTAAGCTGCTCAAGCTGTCACACTCCAAGGGCCAC 480
         |||||||||||| |||||| ||||||||||||||||||||||||||||||||||||||
Db 769 ATTGTCATCCTGTCCTGCTATTATCATCATCATCTCCAAGCTGCTCAAGCTGTCACACTCCAAGGGCCAC 828

Qy 481 CAGAAGGCCAAGGCCCTCAAGACACAGTCAGTCAGTCGCATCCTCTCCAGCTTTCTTGTTCTTGGCCTCTGG 540
         |||||||||||||||||||||||||||| |||||||||||||||||| ||||||||||||||||
Db 829 CAGAAGGCCAAGGCCCTCAAGACACCAGTCATCCTCTCCGGCTTTCTTCTTCTTCGCCTGTTGG 888

Qy 541 CTGCCATATTATGTGGGGATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAA 600
         |||||||||||||||||| ||||| |||||||||||||||||||||||||||||
Db 889 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTTCCTGGAAATCATCAAGCAA 948

Qy 601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGCCCCTCGCCTTC 660
         |||||||||| ||||||||| |||||||||||||| |||||||||||||||| |||||||
Db 949 GGGTGTGAGTTTGAGAACACGTGCACAAGTGGAT-TTCATCACCGAGGGCC--TAGCTTT 1005

Qy 661 TTCCACTGTTGCCTGAACCC 680
         ||| |||||||||||||||
Db 1006 CTTCACTGTTTGTGCTGAACCC 1025

RESULT 10
BU523025
LOCUS   BU523025   906 bp   mRNA   linear   EST 13-SEP-2002
DEFINITION AGENCOURT_10154335 NCI_CGAP_Co24 Mus musculus cDNA clone
            IMAGE:6529663 5', mRNA sequence.
ACCESSION   BU523025
VERSION   BU523025.1 GI:22833475
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 906)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14130 row: c column: 07
            High quality sequence stop: 668.
FEATURES   Location/Qualifiers
     source   1..906
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6529663"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Co24"

http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODF006DE08QP1.

FEATURES
    source   1..1201
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="CS0DF006YJ16"
                    /tissue_type="FETAL BRAIN"
                    /clone_lib="Homo sapiens FETAL BRAIN"
                    /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                    was primed with a NotI-oligo(dT) primer. Five prime end
                    enriched, double-strand cDNA was digested with Not I and
                    cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                    /dev_stage="fetal"
                    vector. Library was not normalized."

ORIGIN

Query Match   65.8%; Score 453.8; DB 13; Length 1201;
Best Local Similarity 84.1%; Pred. No. 2.9e-96;
Matches 573; Conservative 3; Mismatches 85; Indels 20; Gaps 5;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCCTTCTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
         ||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||||
Db 374 CTGCACCTGTCAGTGGCCGACCTCCCTTCTGTCATCACACGCGTTCCCTTCTGGGCAGTTGAT 433

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
         |||| ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db 434 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACA 493

Qy 121 GTCAACCTCTACAGACAGCAGGTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
         |||||||||||||||||||||||| |||||||||||||| |||||||||||||||||||||
Db 494 GTCAACCTCTACAGACAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGTACCTGGCC 553

Qy 181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 554 ATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAACCTGCTGGCTGTGAAAGGTGGTCTAT 613

Qy 241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTTCATCTTTGCCGACGTC 300
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 614 GTTGGGCGTCTGGATCCCGCCCTCCTGCTGACTATTCCGACACTTCTTTGCCAACG-- 671

Qy 301 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACACTCGTGACCGCCTTACCCCGAT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 672 -----------TCAGTRAGGCAGATGACAGATATCTGTGACGCGCTTCTCACCCCAAT 718

Qy 361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCGTGCCCGGC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 719 GACTTGTGGGTGTTGGTGTTCAGTTTCAGCACATCATCATGGTGGGCCTTATCGTGCCTGGT 778

Qy 421 ATCGTCATCCTCCTGTTACTGCATCATCATCATCATCTCAAGCTCATCCT-AGCTTTCTTTGCCTGCTG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 779 ATTGTCATCCTGCCTATTGCATTATCATCTCAAGCTGTCATCCTCAAGCTGTCATCCT-AGCTTTCTTTGCCTGCTG 838

Qy 481 CAGAAGCCAAGGCCCTCAAGACGACAGTCATCCTCATCCTGGCTTCTTCGCCTGTTG 539
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 839 CAGAAGCCAAGGCCCTCAAGACCACAGTCATCCTCATCCTGGCTTCTTCGCCTGTTG 898

Qy 540 GCTGCCATATATGTGGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCA 599
         |||||||||||||||||||||||||||||||||||||||||||| :||| :||||||||
Db 899 GCTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTCCT-GGAAAYATCAAGCA 957

Qy 600 AGGATGTGACTTCGAGAGCATTGTGACAAGTGGATCTCCATCACAGAGGCCTCGCCTTT 659
         ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 958 AGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGAT-TCCATCACCGAGGCC--TAGCTTT 1014

Qy 660 CTTCCACTGTTGCCTGAACCC 680
         | :|| || :||||||||||
Db 1015 TCTYCMCYGTTGTTGTGAACCC 1035

RESULT 12
BX462182

ORIGIN
                    /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                    Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                    Average insert size 1.6 kb. Constructed by Life
                    Technologies. Note: this is a NCI_CGAP Library."

Query Match   67.8%; Score 468; DB 13; Length 906;
Best Local Similarity 97.3%; Pred. No. 1.1e-99;
Matches 498; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCCTTCTGTCATCACACTCCCCTTCTGGGCAGTTGAT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 300 CTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGAT 359

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 360 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 419

Qy 121 GTCAACCTCTACAGACAGCAGGTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 420 GTCAACCTCTACAGACAGCAGGTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCC 479

Qy 181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 480 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTAT 539

Qy 241 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 540 GTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTC 599

Qy 301 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACACTCGTGACCGCCTTTACCCCGAT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 600 AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACACTCGTGACCGCCTTTACCCCGAT 659

Qy 361 AGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGT-GGGTCTCATCCTGCCCGG 419
         |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db 660 AGCCTGTGGATGGTGGTGTTTCAATTCCAGCATATAATGTGGGGTCTCATCCTGCCCGG 719

Qy 420 CATCGTCATCCTCCTGTTACTGCATCATCATCATCTCTAAGCTCACTCTCAAGGG--- 476
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 720 CATCGTCATCCTCCTGTTACTGCATCATCATCATCTCTAAGCTGTCACACTCCAAGGGGGCC 779

Qy 477 CCACCAGAAGCGCAAGGCCCTCAAGACGACAGACAG 508
         |||||||| || ||||||||||||||||||||||||
Db 780 ACCCAGAAAACGCCAGGCCCCTCAAGACGACAG 811

RESULT 11
BX418530
LOCUS   BX418530   1201 bp   mRNA   linear   EST 13-MAY-2003
DEFINITION   BX418530 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
             CS0DF006YJ16 5-PRIME Homo sapiens cDNA sequence.
ACCESSION   BX418530
VERSION   BX418530.1 GI:30658606
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 1272.r For
             more information about this cluster, see
             http://www.genoscope.cns.fr/
             cgi-bin/cluster.cgi?seq=CSODF006DE08QP1&cluster=1272.r Contact :
             Feng Liang Email : fliang@lifetech.com URL :

```
LOCUS       BX462182                1201 bp    mRNA    linear  EST 22-MAY-2003
DEFINITION  BX462182 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG003YO19 5-PRIME, mRNA sequence.
ACCESSION   BX462182
VERSION     BX462182.1  GI:31031389
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DG003AH10QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DG003AH10QP1
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG003YO19"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

Query Match          63.3%;  Score 436.8;  DB 13;  Length 1201;
Best Local Similarity  82.8%;  Pred. No. 3e-92;
Matches  510;  Conservative  13;  Mismatches  77;  Indels  16;  Gaps  2;

Qy      1 CTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCCCTTCTGGGCAGTTGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    357 CTGCACCTGTCAGTGGCCGGNCCTCCTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT 416

Qy     61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCGTGTCCATATCATCTCACACT 120
          || |||||||||||||||||||||||||| ||||| |||||||||||||||||||||||||
Db    417 GCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCCAAGGCAGTCCATGTCATCTACACA 476

Qy    121 GTCAACCTTCTACAGCAGCCGGTTCTCATCCTGGCCTTCATCGACCGGTACCTCGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 NTCAACCTTCTACAGCAGTGTCCCATCCTGCCCTTCATCCTGGCCTTCATCGACCGGTACCTCGCC 536

Qy    181 ATTGTCCACGGCACCACCACAGTCAAAGGCCAAGGAAACTGCTGGCGCAAAAGGCAGTCTAT 240
          || ||||||||||||||||||||||| |||||||| || ||||||||||||||||||||
Db    537 ATCGTCCACGGCACCACCACAGTCAAACAGTGCAGAGGGCCAAGGAAGCGTGTTGGCTGCAAAAAGGKGGKCTAT 596

Qy    241 GTGGGGCGTCTGGATCCCAGCCCTCCTCGACTATACCTGACTTCATCTTTGCCGACGTC 300
          ||||||| ||||||||||||||||||||||||| |||||||||||||||||
Db    597 GTTGGGCGTCTGGGATCCCGCCCTCCTGCTGACGTACTACTTTCCCGACTTCTGTCATCTTTGCCAACG-- 654

Qy    301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTACCCGAT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    655 --------TCAGTAAGGCAGGATGACAGATATCTGTGAMCGCTTCTACCCCCAAT 701

Qy    361 AGCCTGTGGATGGTGGTGTTCAATTCAGCATATAATGTGGGGTCTCATCCTGCCCGGC 420
          |||| ||| ||||||||||| |||| ||||||||||||||||| |||||||||||
Db    702 GACTTGTGGGKGGKTGTGTTCCAGTTYMAGCAMATMATGTTGTTGGCCTGATGCCTTATCCTGCCCTGGT 761
```

RESULT 13
```
LOCUS       CD466324                 749 bp    mRNA    linear  EST 04-JUN-2003
DEFINITION  LeukoN2_3_D11.g1_A024 Unstimulated peripheral blood leukocytes N2
            Equus caballus cDNA clone LeukoN2_3_D11_A024 5', mRNA sequence.
ACCESSION   CD466324
VERSION     CD466324.1  GI:31387592
KEYWORDS    EST.
SOURCE      Equus caballus (horse)
ORGANISM    Equus caballus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE   1  (bases 1 to 749)
AUTHORS     Vandenplas,M., Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S.,
            Moore,J.N., Liang,C., Sun,F., Sullivan,R., Shah,M., and Pratt,L.H.
TITLE       An EST database from equine (Equus caballus) unstimulated
            peripheral blood leukocytes
JOURNAL     Unpublished (2003)
COMMENT     Other ESTs: LeukoN2_3_D11.b1_A024
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; tissue and RNA were prepared in the Department of Large
            Animal Medicine, University of Georgia; sequencing done in the
            Laboratory for Genomics and Bioinformatics, University of Georgia.
            Sequence ends have been trimmed to exclude vector and regions below
            Phred quality 16. Three-prime sequences are presented as their
            reverse complement and have been trimmed to exclude polyA.
            Seq primer: Sug5 (CTTCTGCTCTAAAAGCTGCG)
FEATURES             Location/Qualifiers
     source          1..749
                     /organism="Equus caballus"
                     /mol_type="mRNA"
                     /strain="thoroughbred"
                     /db_xref="taxon:9796"
                     /clone="LeukoN2_3_D11_A024"
                     /sex="male"
                     /tissue_type="blood"
                     /cell_type="leukocytes"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Unstimulated peripheral blood leukocytes N2"
                     /note="Organ: circulatory system; Vector: pME18S-FL3;
                     Site_1: XhoI; Site_2: XhoI; The library was prepared from
                     polyA+ RNA from unstimulated equine peripheral blood
                     leukocytes isolated from a healthy adult horse.
                     Double-stranded cDNA was cloned unidirectionally into
                     different DraIII sites of the pME18S-FL3 vector (5-prime
                     DraIII site is CACTGTGTG, 3-prime DraIII site is
                     CACCATGTG). XhoI excises the cDNA insert."
```

```
Qy    421 ATGTGTCATCCTCCTGCTTACTGCGATCATCATCTTAAGCTGTCACACTCCAGGGGCCAC 480
          ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db    762 ATTGTCATCCTGTCCTGCTATTGCATTATCATCTTCCAAGCTGTCACACTCCAAGGGSCAC 821

Qy    481 CAGAAGGCGAAGGCCCTCAAGAGACGACAGTCATCATCCTCAGCTTTCTTGCCTGCTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db    822 CAGAAGGCGAAGGCCCTCAAGACCATCATCATCCTCCTGCTTTCTTCCTGGCCGTGTTGG 881

Qy    541 CTGCCATATATGTGGGATCAGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAA 600
          ||||||:||:|||||||||||| ||||||||||||||||:|||||:||||||||||||||||
Db    882 CTGCCTTACTACATTGGGATMA-SATCGACTCCTTCATCCTCCTGGAAATCATMAAGCAA 940

Qy    601 GGATGTGACTTCGAGA 616
          ||| :|||||:|:|||
Db    941 GGGKGTGAGTTTRARA 956
```

ORIGIN

Query Match   63.1%;   Score 435.6;   DB 14;   Length 749;
Best Local Similarity   83.2%;   Pred. No. 4.5e-92;
Matches   514;   Conservative   0;   Mismatches   89;   Indels   15;   Gaps   1;

```
Qy     73  TGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTAC  132
           ||||||||||||| ||:|||||||||||||| |||:||||:|||:|| ||||||||||
Db      1  TGGTACTTTGGGAAGTTCCTGTGCAAGGCCGTTCATGTCATCACAGTCAACCTCTAC    60

Qy    133  AGCAGGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGCCATTGTCCACGCC  192
           ||||||| |||||||||||||||||||||||  |||| ||||||||||||||||||||
Db     61  AGCAGTCGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGTACCTGGCTATCGTCCACGCC  120

Qy    193  ACCAACAGTCAAAGGCCAAGGAAACTGTCGCTGCTGAAAAGCAGTCTATGTGGGCGTCTGG  252
           ||||| ||:|||||||||||||| ||| ||||||| ||||||| ||| |||||||:|||||
Db    121  ACCAACAGTCAGAAGGCCAAGGAAGCTGTTGGCTGATGAAAAAGTGGTCTATGTGGTCTGG  180

Qy    253  ATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGAC  312
           |||||||||| |||| ||||   ||||||||||||||||||||| ||||| ||||
Db    181  ATACCTGCTCTTCTGCTGACTATTCCCACACTTCATCTT TGCCAACG--------------  226

Qy    313  ATCAGTCAGGGGGATGACAGTACATCGTGACCGGCCTTTACCCGACTAGCCTGTGGATG  372
           |||:||||||||||||||||| || |||: |||| |||||| |||||:||||||||||
Db    227  -TCAGGAGGGGGGATGGGAGGTATATCGTGACCGGTTCTCACCCCAGTGACTTGTGGTTG  285

Qy    373  GTGGTGTTCAATTCCAGCATATAAATGGTGGGGTCTCATCTCGCCCGGCATCGTCATCCTC  432
           ||||||:||||||| ||||:|| |||||:||||||||||||||| || ||||||||||||
Db    286  GTGGTGTTCCAGTTCAGCACATAATGGTGTCGGGCCTTATCCTGCCAGGCATCGTCATCCTA  345

Qy    433  TCCTGTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCCGCAAG  492
           |||||||||:|| ||||||||||||||:|||||||||||||||||||| ||||||||||||||
Db    346  TCCTGCTATTGCGTTATCATCATCCCAAGCTCTCCCACTCCAAGGGCTACCAGAAAGCCAAG  405

Qy    493  GCCCTCAAGACGACAGTCATCCTCCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATTAT  552
           ||||| ||| ||||||||||||||||||||||||||:||||:||||||||||||||||||
Db    406  GCCCTAAAGACCACAGTTATCCTCCATCCTAGCTTTCGTTCGCCTGCTGGCTCCCTACTAC  465

Qy    553  GTGGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTC  612
           ||||| |||||||||||||||||| |||||||| |||||||||||||||||||||||||||
Db    466  ATTGGGGATCAGCATCGACTCCTTCATCCTCCTGGAAATCATCAGCGAGGATGTGAGTTT  525

Qy    613  GAGAGCATTGTGCACAAGTGGATCCTCCATCACAGAGAGGCCCTTCTTCCCACTGTTGC  672
           |||||||||||||| ||||||||||||||||||||||||| |||||||||||||||||||
Db    526  GAGAGCACTGTGCACAAGTGGATTTCATCACCGAGAGGCCCTTCTTTTCCACTGTTGC  585

Qy    673  CTGAACCCCATCCTCTAT  690
           ||||| |||||||||||
Db    586  CTGAATCCCATCCTCTAT  603
```

RESULT 14
BG145042
LOCUS   BG145042   582 bp   mRNA   linear   EST 01-FEB-2001
DEFINITION   ut75a01.y1 Soares_mouse_NMGB_bcell Mus musculus cDNA clone
             IMAGE:3333409 5' similar to SW:CCR4_MOUSE P70658 C-X-C CHEMOKINE
             RECEPTOR TYPE 4 ;, mRNA sequence.
ACCESSION   BG145042
VERSION   BG145042.1   GI:12648446
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 582)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
  JOURNAL   Unpublished (1997)
  COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1077573
Seq primer: -40RP from Gibco
High quality sequence stop: 503.

ORIGIN

Query Match   63.0%;   Score 435;   DB 12;   Length 582;
Best Local Similarity   100.0%;   Pred. No. 5.6e-92;
Matches   435;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    256  CCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATC  315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  CCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATC   60

Qy    316  AGTCAGGGGGATGACAGTACATCGTGACCGGCCTTACCCGATAGCCTGTGGATGGTG  375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  AGTCAGGGGGATGACAGTACATCGTGACCGGCCTTACCCGATAGCCTGTGGATGGTG  120

Qy    376  GTGTTTCAATTCCAGCATATAAATGGTGGGGTCTCATCTCGCCCGGCATCGTCATCCTCTCC  435
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GTGTTTCAATTCCAGCATATAAATGGTGGGGTCTCATCTCGCCCGGCATCGTCATCCTCTCC  180

Qy    436  TGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCC  495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCC  240

Qy    496  CTCAAGACGACAGTCATCCTCCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTG  555
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CTCAAGACGACAGTCATCCTCCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTG  300

Qy    556  GGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGGATGGACTTCGAG  615
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  GGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGGATGGACTTCGAG  360

Qy    616  AGCATTGTGCACAAGTGGATCCTCCATCACAGAGAGGCCCTTCTTCCCACTGTTGCCTG  675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AGCATTGTGCACAAGTGGATCCTCCATCACAGAGAGGCCCTTCTTCCCACTGTTGCCTG  420

Qy    676  AACCCCATCCTCTAT  690
           |||||||||||||||
Db    421  AACCCCATCCTCTAT  435
```

RESULT 15
BX362477
LOCUS   BX362477   952 bp   mRNA   linear   EST 05-MAY-2003
DEFINITION   BX362477 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
             Homo sapiens cDNA clone CS0DJ01SYP09 5-PRIME, mRNA sequence.
ACCESSION   BX362477
VERSION   BX362477.1   GI:30380640
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 952)

AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DJ015CH05QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DJ015CH05QP1.

FEATURES   Location/Qualifiers
     source   1..952
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ015YP09"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with the Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Db   895 CTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCYCTKGSAA 942

Search completed: May 17, 2004, 16:09:01
Job time : 1999.6 secs

Query Match   62.6%;   Score 431.8; DB 13; Length 952;
Best Local Similarity   84.5%;   Pred. No. 4e-91;
Matches 497; Conservative   4; Mismatches 72; Indels 15; Gaps   1;

Qy   1 CTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACTCCCTTCTGGCAGTTGAT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   370 CTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCTTCCCTTCTGGGCAGTTGAT   429

Qy   61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   430 GCCGTGGCAAACTGGTACTTTGGGAACTTCCCATGCAAGGCAGTCCATGTCATCTACACA   489

Qy   121 GTCAACCTTACAGCAGCGTTCATCTGGCCTTCATCGGCCTTGACCGGTACTTCGCC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   490 GTCAACCTCTACGGCAGTGTCCTCTCCATCCTGGCCTTCATCAGTCTGGACCGCTACCTGCC   549

Qy   181 ATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGAAAAGGCAGTCTAT   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   550 ATTGTCCACGCCACCAACAGTCAAGGAGCCGAGGCAAGGAAGCTGTTGGCTGAAAAGGTGGTTCTAT   609

Qy   241 GTGGGCGTTCTGGATCCCAGCCCTTCCTCGACTATAACCTGACTTCATCTTTGCCGACGTC   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   610 GTTGGCGTCTGGAATCCCTGCCCTTCCTCGCTGACTATTCCCGACTTCATCTTTGCCAACG--   667

Qy   301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGTACATCGTGACCGCCTTACCCCGAT   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 --------TCAGTANGGCAGATGACAGATATCTGTGACCGCTTCACCCCAAT   714

Qy   361 AGCCTGTGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTTCTCATCCTGCCCGGC   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   715 GACTTGTGGGTGGGTTGTGTTCCAGTTTGCAGCCACATCATGGTTGGCCTTATCCTGCCCTGGT   774

Qy   421 ATCGTCATCCTCTCCTCCGTTACTGGCATCATCATCCTAGCTGTCCACACTCCAAGGGGCCAC   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   775 ATTGTCATCCTCGTCTGCTATTGCATTATCATCTCCAAGGTGTCACACTCCCAAGGGGCCAC   834

Qy   481 CAGAAGCGCAAGGCCCTAAGACGACAGCAGTCATCCATCATCTTCTTTGCCTGCTGG   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   835 CAGAAGCGCAAGGCCCTAAGACCAACAGTCATCCATCCTCGCTTTCTTCTTCGCCTGTTGG   894

Qy   541 CTGCCATTATGTGGGGATCAGGCATCGACTCCTTCATCCTTTTGGGA   588
           |||| || |||||||||||||| || ||| ||||| |:|: |.|:: |

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: May 17, 2004, 04:56:59 ; Search time 7592.63 Seconds
                                         (without alignments)
                                         10714.982 Million cell updates/sec

Title: US-09-367-052-1
Perfect score: 1877
Sequence: 1877
Sequence: 1 ccatcctaatacgactcact.........aaaaaaaaaaaaaaaaaaaa 1877
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21571516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
     1: gb_ba:*
     2: gb_htg:*
     3: gb_in:*
     4: gb_om:*
     5: gb_ov:*
     6: gb_pat:*
     7: gb_ph:*
     8: gb_pl:*
     9: gb_pr:*
    10: gb_ro:*
    11: gb_sts:*
    12: gb_sy:*
    13: gb_un:*
    14: gb_vi:*
    15: em_ba:*
    16: em_fun:*
    17: em_hum:*
    18: em_in:*
    19: em_mu:*
    20: em_om:*
    21: em_or:*
    22: em_ov:*
    23: em_pat:*
    24: em_ph:*
    25: em_pl:*
    26: em_ro:*
    27: em_sts:*
    28: em_un:*
    29: em_vi:*
    30: em_htg_hum:*
    31: em_htg_inv:*
    32: em_htg_other:*
    33: em_htg_mus:*
    34: em_htg_pln:*
    35: em_htg_rod:*
    36: em_htg_mam:*
    37: em_htg_vrt:*
    38: em_sy:*
    39: em_htgo_hum:*
    40: em_htgo_mus:*
    41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1877	100.0	1877	10	D87747	D87747 Mus musculu
2	1797.4	95.8	1817	10	BC031665	BC031665 Mus muscu
3	1683.6	89.7	1809	10	AB000803	AB000803 Mouse mRN
4	1602.6	85.4	175370	2	AC145861	AC145861 Pan trogl
5	1602.6	85.4	165215	2	AC147556	AC147556 Mus muscu
6	1474.8	78.6	1575	10	MMLESTRPT	X99582 M.musculus
7	1384.6	73.8	3770	10	MMLESTRGN	X99581 M.musculus
8	1216.6	64.8	1223	10	MMU59760	U59760 Mus musculu
9	1169.4	62.3	190149	2	AC122097	AC122097 Rattus no
10	1152.6	61.4	1422	10	MMLCR13	Z80113 Mus musculu
11	1143.2	60.9	1180	10	MMLCR12	Z80112 Mus musculu
12	1114	59.4	240931	2	AC107097	AC107097 Rattus no
13	1055	56.2	3356	10	MMU65580	U65580 Mus musculu
14	974.4	51.9	1737	6	AR015970	AR015970 Sequence
15	974.4	51.9	1737	6	AR060748	AR060748 Sequence
16	974.4	51.9	1737	6	AR070433	AR070433 Sequence
17	974.4	51.9	1737	6	AR103430	AR103430 Sequence
18	974.4	51.9	1737	6	I13753	I13753 Sequence 4
19	974.4	51.9	1737	6	I24455	I24455 Sequence 3
20	960.4	51.2	1662	6	BC020968	BC020968 Homo sapi
21	958.8	51.1	1362	10	RNU54791	U54791 Rattus norv
22	952.8	50.8	1667	9	AK129916	AK129916 Homo sapi
23	952.8	50.8	1679	6	AX287116	AX287116 Sequence
24	952.8	50.8	1679	6	AX548790	AX548790 Sequence
25	952.8	50.8	1679	6	AX705078	AX705078 Sequence
26	951.8	50.7	1664	6	E37853	E37853 Anti-HIV ag
27	951.8	50.7	1664	6	AR222504	AR222504 Sequence
28	951.8	50.7	1664	6	AR380668	AR380668 Sequence
29	951.8	50.7	1664	9	HUMHM89	D10924 Homo sapien
30	951.8	50.7	1664	11	G28514	G28514 SHGC-31461
31	947.6	50.5	1645	9	HSNPYRLA	X71635 H.sapiens m
32	944.8	50.3	1670	6	AX409496	AX409496 Sequence
33	944.8	50.3	1670	6	AX63641	AX63641 Sequence
34	944.8	50.3	1670	9	HUMGPCR	L06797 Human (clon
35	939.8	50.1	1944	9	E33806	E33806 Human splic
36	938.2	50.0	1950	9	AF147204	AF147204 Homo sapi
37	934	49.8	3733	9	HSCXCR4	Y14739 Homo sapien
38	934	49.8	5161	9	AF005058	AF005058 Homo sapi
39	934	49.8	172281	2	AC068492	AC068492 Homo sapi
40	933.6	49.7	1582	6	AX481810	AX481810 Sequence
41	932.4	49.7	8747	9	AF052572	AF052572 Homo sapi
42	932.4	49.7	8747	9	HSA224869	AJ224869 Homo sapi
43	921.2	49.1	1637	9	HUMSTSR	M99293 Homo sapien
44	908.8	48.4	1050	10	AF452185	AF452185 Rattus no
45	907.4	48.3	1689	4	FCA9816	AJ009816 Felis cat

ALIGNMENTS

RESULT 1
D87747
D87747
LOCUS   D87747   1877 bp   mRNA   linear   ROD 07-FEB-1999
DEFINITION   Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION   D87747
VERSION   D87747.1 GI:1772445
KEYWORDS   murine CXCR-4.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (sites)
AUTHORS   Nagasawa,T., Nakajima,T., Tachibana,K., Iizasa,H., Bleul,C.C.,
             Yoshie,O., Matsushima,K., Yoshida,N., Springer,T.A. and
             Kishimoto,T.

TITLE   Molecular cloning and characterization of a murine pre-B-cell
           growth-stimulating factor/stromal cell-derived factor 1 receptor, a
           murine homolog of the human immunodeficiency virus 1 entry
           coreceptor fusin
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)
MEDLINE   97121456
PUBMED   8962122
REFERENCE 2 (bases 1 to 1877)
AUTHORS   Nagasawa,T.T.
TITLE   Molecular cloning of murine PBSF/SDF-1 receptor
JOURNAL   Unpublished
REFERENCE 3 (bases 1 to 1877)
AUTHORS   Nagasawa,T.
TITLE   Direct Submission
JOURNAL   Submitted (05-SEP-1996) Takashi Nagasawa, Research Institute, Osaka
           Medical Center, Department of Immunology; Murodocho 840, Izumi,
           Osaka 590-02, Japan (E-mail:immunol@osk.threewebnet.or.jp,
           Tel:0725-56-1220, Fax:0725-57-3021)
FEATURES   Location/Qualifiers
     source   1..1877
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="DW34"
                     /cell_type="pre-B cell"
                     /tissue_type="bone marrow"
     gene   1..1877
                     /gene="murine CXCR-4"
     CDS   120..1199
                     /gene="murine CXCR-4"
                     /function="PBSF/SDF-1 receptor"
                     /codon_start=1
                     /evidence=experimental
                     /product="murine CXCR-4"
                     /protein_id="BAA13451.1"
                     /db_xref="GI:1772446"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGGDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILLLAFFACWLPYVVGISDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"

ORIGIN
Query Match   100.0%;   Score 1877;   DB 10;   Length 1877;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1877;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 CCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGTGCAGGTAGCAG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGTGCAGGTAGCAG   60

Qy   61 TGACCCCTCGAGGGGTTTGGCTCCGGTAACCACCACCACGGCTGTAGAGGCGAGTGTTGCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGACCCCTCGAGGGGTTTGGCTCCGGTAACCACCACCACGGCTGTAGAGGCGAGTGTTGCCA 120

Qy   121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTTCTGATACTACTTCTGAAGAAGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTTCTGATACTACTTCTGAAGAAGA 180

Qy   181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGA 240

Qy   241 TCTTCCTGCCCACCATCACTTCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCTTCCTGCCCACCATCACTTCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 300

Qy   301 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 360

Qy   361 ACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCTTCGGCCAGTGATGCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCTTCGGCCAGTGATGCCA 420

Qy   421 TGGCTGACTGGTACTTTGGGAAAATTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TGGCTGACTGGTACTTTGGGAAAATTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480

Qy   481 ACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGGCCATTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGGCCATTG 540

Qy   541 TCCACGCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCCACGCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600

Qy   601 GCGTCTGGATCCCAGCCCTCCTCGCCCGCTATACCTGACTTCATCTTTGCCGACGTCAGCC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCGTCTGGATCCCAGCCCTCCTCGCCCGCTATACCTGACTTCATCTTTGCCGACGTCAGCC 660

Qy   661 AGGGGACATCAGTCAGGGGATGACAGAGTACATCTGTGACCGCCTTACCCGATAGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AGGGGACATCAGTCAGGGGATGACAGAGTACATCTGTGACCGCCTTACCCGATAGCC 720

Qy   721 TGTGGATGGTGGTGTTCAATTCCAGCATATAAATGGTGGGTCTTCATCCTGCCCGGCATCG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TGTGGATGGTGGTGTTCAATTCCAGCATATAAATGGTGGGTCTTCATCCTGCCCGGCATCG 780

Qy   781 TCATCCTCTGTGTTACTGCCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TCATCCTCTGTGTTACTGCCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGA 840

Qy   841 AGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTCTTCTTGCTGGGCTGC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTCTTCTTGCTGGGCTGC 900

Qy   901 CATATTATGTGGGGATCAGCATGCATCGACTCCTTTGGGAGTCATCAAGCAAGGAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CATATTATGTGGGGATCAGCATGCATCGACTCCTTTGGGAGTCATCAAGCAAGGAT 960

Qy   961 GTGACTTGGAGAGACCATTGTGCAAAGTGGGATCTCCATCACAGAGGCCCTGCCTTCTTCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GTGACTTGGAGAGACCATTGTGCAAAGTGGGATCTCCATCACAGAGGCCCTGCCTTCTTCC 1020

Qy   1021 ACTGTGCCTGAACCCCATCCTCTATGGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACTGTGCCTGAACCCCATCCTCTATGGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCC 1080

Qy   1081 AGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGC 1140

Qy   1141 GGGGTGGACACTCTTCCGTCTCCACGAGTCAGAATCCTCCAGTTTTCACTCAGCTAAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGGGTGGACACTCTTCCGTCTCCACGAGTCAGAATCCTCCAGTTTTCACTCAGCTAAC 1200

Qy   1201 CCTATGCAAAGACTTATATAATATATATATATA'TTTTATGTTAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCTATGCAAAGACTTATATAATATATATATA'TTTTATGTTAC 1260

Qy   1261 ACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTAAATTGA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTAAATTGA 1320

Qy   1321 CTGTTGGGAGTTTATGTTCCTCTTAGTTTTGTGAGGTTTGACTAATTTATATAAAATATT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGTTGGGAGTTTATGTTCCTCTTAGTTTTGTGAGGTTTGACTAATTTATATAAAATATT 1380

Qy   1381 GTTTTTTGTTTGTTTGTTCATGTGAATGAGTGAATGGCGTCTAGGCAAGTTCTTAGT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTTTTGTTTGTTTGTTCATGTGAATGAGTGAATGGCGTCTAGGCAAGTTCTTAGT 1440

Qy   1441 AGCTGTTTATCTGTGTGTAGGACTGTGAACTGTAGAGGAAGAAACTGAACATTCCAGAA 1500

```
REMARK      Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
COMMENT     USA
            NIH-MGC Project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@cm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 61 Row: d Column: 24
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.
FEATURES             Location/Qualifiers
     source          1..1817
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:36266 IMAGE:3592479"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1..1817
                     /gene="Cxcr4"
                     /note="synonyms: LESTR, fusin, PB-CKR, PBSF/SDF-1, CD184"
                     /db_xref="LocusID:12767"
                     /db_xref="MGI:109563"
     CDS             78..1157
                     /codon_start=1
                     /product="chemokine (C-X-C motif) receptor 4"
                     /protein_id="AAH31665.1"
                     /db_xref="GI:21618782"
                     /db_xref="LocusID:12767"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQDISQGDDRYICDRLYPDSLWMVFQFQHIMVGLLLP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"
     misc_feature    246..1004
                     /note="7tm_1; Region: 7 transmembrane receptor (rhodopsin
                     family)"
                     /db_xref="CDD:pfam00001"
ORIGIN
Query Match       95.8%;     Score 1797.4;     DB 10;   Length 1817;
Best Local Similarity    99.9%;   Pred. No. 0;
Matches 1809;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy         50 GCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCGTGTAGAGC 109
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 GCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCGTGTAGAGC 67

Qy        110 GAGTGTTGCCATGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GAGTGTTGCCATGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 127

Qy        170 AGTGGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1441 AGCTGTTTATCTGTGTGTAGGACTGTAGGAACTGTAGAGGAAGAAACTGAACATTCCAGAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1501 TGTGTGGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTGCTGCGCATAATCTCTTCA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TGTGTGGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTGCTGCGCATAATCTCTTCA 1560
Qy       1561 TTCCGAGGAGCACCCCCACCCCCACCCCCACCCCCACCCCCATTCTTAAATTGTTTGGTTAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTCCGAGGAGCACCCCCACCCCCACCCCCACCCCCACCCCCATTCTTAAATTGTTTGGTTAT 1620
Qy       1621 GCTGTGTGATGGTTGTTTGTTTGTGTTTTTTTTTTTTTCTGTAAA 1680
              |||||||||||||||||||||||||||||||||||
Db       1621 GCTGTGTGATGGTTGTTTGTTTGTGTTTTTTTTTTTTTTTTCTGTAAA 1680
Qy       1681 AGATGGCACTTAAAACCAAAGCCTGAAAATGGTGTAGAAAATGCTGGGGTTTTTTTTGTTT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGATGGCACTTAAAACCAAAGCCTGAAAATGGTGTAGAAAATGCTGGGGTTTTTTTTGTTT 1740
Qy       1741 GTTTGTTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GTTTGTTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTG 1800
Qy       1801 TATTACATTGTTAATAAAAGTCAATGATGAATAACTTAAAAAAAAAAAAAAAAAAAAAAAA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TATTACATTGTTAATAAAAGTCAATGATGAATAACTTAAAAAAAAAAAAAAAAAAAAAAAA 1860
Qy       1861 AAAAAAAAAAAAAAA 1877
              |||||||||||
Db       1861 AAAAAAAAAAAAAAA 1877

RESULT 2
BC031665
LOCUS       BC031665              1817 bp    mRNA      linear   ROD 07-OCT-2003
DEFINITION  Mus musculus chemokine (C-X-C motif) receptor 4, mRNA (cDNA clone
            MGC:36266 IMAGE:3592479), complete cds.
ACCESSION   BC031665
VERSION     BC031665.1  GI:21618781
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1817)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1817)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
```

```
Db   128  AGTGGGGTCTCGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAGCTCCA  187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   230  TTTCAATAGGATCTTCCTGCGCCACCATCTACTTCATCATCTTCTTGACTCGGCATAGTCGG  289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188  TTTCAATAGGATCTTCCTGCGCCACCATCTACTTCATCATCTTCTTGACTCGGCATAGTCGG  247

Qy   290  CAATGGATTGGTGATCTCGGTCATGGGTTACCAGAGAAGCTAAGGACGATGACGACAA  349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248  CAATGGATTGGTGATCTCGGTCATGGGTTACCAGAGAAGCTAAGGACGATGACGACAA  307

Qy   350  GTACCGGCTGCACCTGTCAGTGGCTGACACTCCCCTTCTCCCTTCTGGGC  409
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308  GTACCGGCTGCACCTGTCAGTGGCTGACACTCCCCTTCTCCCTTCTGGGC  367

Qy   410  AGTTGATGCCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT  469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368  AGTTGATGCCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT  427

Qy   470  CTACACTGTCAACCTCACAGCAGGCGTTCTCATCTCGGCCTTCATCAGCCTGGACCGGTA  529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428  CTACACTGTCAACCTCACAGCAGGCGTTCTCATCTCGGCCTTCATCAGCCTGGACCGGTA  487

Qy   530  CCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGC  589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488  CCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGC  547

Qy   590  AGTCTATGTGGGGGTCTGGATCCCAGCCCTCCTCCGACTATACCTGACTTCATCTTTGC  649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548  AGTCTATGTGGGGGTCTGGATCCCAGCCCTCCTCCGACTATACCTGACTTCATCTTTGC  607

Qy   650  CGACGTCAGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCGCCTTTA  709
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608  CGACGTCAGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCGCCTTTA  667

Qy   710  CCCGATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGTGGGTCTCATCCT  769
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668  CCCGATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGTGGGTCTCATCCT  727

Qy   770  GCCCGGGCATCGTCATCCTCCTGTTACTGACTCATCATCATCTCTAAGCTGTCACACTCCAA  829
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728  GCCCGGGCATCGTCATCCTCCTGTTACTGACTCATCATCATCTCTAAGCTGTCACACTCCAA  787

Qy   830  GGGCCACCAGAAGCGCAAGGCCCTCAAGACAGCGCAAGGCCCTCAAGACAGCAGTCTTCTTGC  889
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788  GGGCCACCAGAAGCGCAAGGCCCTCAAGACAGCGCAAGGCCCTCAAGACAGCAGTCTTCTTGC  847

Qy   890  CTGCTCGGCTGCCATATATGTGGGGATCAGCATCAGCATCAGCATCCTCCTTCATCCTTTTGGGAGTCAT  949
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848  CTGCTCGGCTGCCATATATGTGGGGATCAGCATCAGCATCAGCATCCTCCTTCATCCTTTTGGGAGTCAT  907

Qy   950  CAAGCAAGGATGATGACTTCGAGAGCAATTGTGCACAAGTGGATCTCCATCACAGAGGCCCT  1009
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908  CAAGCAAGGATGATGACTTCGAGAGCAATTGTGCACAAGTGGATCTCCATCACAGAGGCCCT  967

Qy  1010  CGCCTTCTTCCACTGTTGCCTGAACCCATCCTCTATGCCTTCCTCGGGGCCAAGTTCAA  1069
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968  CGCCTTCTTCCACTGTTGCCTGAACCCATCCTCTATGCCTTCCTCGGGGCCAAGTTCA  1027

Qy  1070  AAGCTCTGCCCAGCATGCACTCAACTCAACTCCATGAGGACGAGAGGCTCCAGCCTTCAAGATCCTTTC  1129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1028  AAGCTCTGCCCAGCATGCACTCAACTCAACTCCATGAGGACGAGAGGCTCCAGCCTTCAAGATCCTTTC  1087

Qy  1130  CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGGAATCCTCCAGTTTTCA  1189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088  CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGGAATCCTCCAGTTTTCA  1147

Qy  1190  CTCCAGCTAACCCTTATGCAAAGACTTATATATATATAGATAAAGAACTT  1249
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148  CTCCAGCTAACCCTTATGCAAAGACTTATATATATATAGATAAAGAACCT  1207

Qy  1250  TTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTT  1309
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208  TTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTT  1267
```

```
Qy  1310  TTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGACTTAATTT  1369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268  TTTTTAATTGACTGTTGGGAGTTTATGTTCCCTCTAGTTTTGTGAGGTTTGACTTAATTT  1327

Qy  1370  ATATAAAATATTGTTTTGTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGGCC  1429
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328  ATATAAAATATTGTTTTGTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGGCC  1387

Qy  1430  AAGTTCTTAGTAGCTGTTTATCTGTGTGAGGACTGTAGAACTGTAGAGGAAGAAACTGA  1489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388  AAGTTCTTAGTAGCTGTTTATCTGTGTGAGGACTGTAGAAACTGTAGAGGAAGAAACTGA  1447

Qy  1490  ACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTGCAGCTGTTGCTGCA  1549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1448  ACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTGCAGCTGTTGCTGCA  1507

Qy  1550  TAATCTCTTCATTCCGAGAGCACCCACCCCACCCCCACCCCATTCTTAAAT  1609
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1508  TAATCTCTTCATTCCGAGAGCACCCACCCCACCCCCACCCCATTCTTAAAT  1567

Qy  1610  TGTTTGGTTATGCTGTGTGATGGTTGTTTGTTGTTTTTTGTTTGTGTTGTTTTT  1669
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1568  TGTTTGGTTATGCTGTGTGATGGTTGTTTGTTGTTTTTTGTTTGTTG-TTTTT  1626

Qy  1670  TTTTCTGTAAAAGATGGCACTTAAAACCAAAGCTCGAAATGGTGGTGTAGAAATGCTGGGGT  1729
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1627  TTTTCTGTAAAAGATGGCACTTAAAACCAAAGCTCGAAATGGTGTGTAGAAATGCTGGGGT  1686

Qy  1730  TTTTTTGTTGTTTGTTTGTTTTCAGTTTTCAGTTTTCAAGAGTAGATTCAGTTCAGTTCCCTACAAAT  1789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1687  TTTTTTGTTGTTTGTTTGTTTTCAGTTTTCAGTTTTCAAGAGTAGATTCAGTTCAGTTCCCTACAAAT  1746

Qy  1790  GTACAGTCTTGTATTACATTGTTAATAAAAGTCAATGATAAACTTAAAAAAAAAAAA  1849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1747  GTACAGTCTTGTATTACATTGTTAATAAAAGTCAATGATAAAACTTAAAAAAAAAAA  1806

Qy  1850  AAAAAAAAA  1860
          ||||||||
Db  1807  AAAAAAAAA  1817
```

```
CDS          /dev_stage="E17"
             72. -1151
             /note="fusin"
             /codon_start=1
             /product="murine CXCR-4"
             /protein_id="BAA19187.1"
             /db_xref="GI:1816446"
             /translation="MEPISVSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLP
             TIYFIIFLTGVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
             DWYFGKFLCKAVHIIYTVNLYSSVLLIAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
             GVWIPALLLTIPDFIFADVSOGDISQGDDRYICDRLYPDSLWMVFQFQHIMVGLVLP
             GIVILSCYCIIISKLSHSKGHQKRKALKTTVILLIAFFACWLPYYVGISIDSFILLGV
             IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMRGSSLK
             ILSKGKRGGHSSVSTESESSSFHSS"

ORIGIN

Query Match      89.7%;   Score 1683.6;   DB 10;   Length 1809;
Best Local Similarity  97.7%;   Pred. No. 0;
Matches 1765;   Conservative  0;   Mismatches  24;   Indels 17;   Gaps 5;
```

```
Qy    50 GCAGGTAGCAGTG-ACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAG  108
         ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     1 GCAGGTAGCAGTGAAACCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAG   60

Qy   109 CGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAG 168
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAG 120

Qy   169 AAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCGCTTCCGGATGAAAACGTCC   228
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCGCTTCCGGATGAAAACGTCC   180

Qy   229 ATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCG 288
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCG 240

Qy   289 GCAATGGATGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACA  348
         ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACA 300

Qy   349 AGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGG  408
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGG  360

Qy   409 CAGTTGATGCCATGGCCATGGCTACTGTGGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCA 468
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CAGTTGATGCCATGGCCATGGCTACTGTGGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCA 420

Qy   469 TCTACACTGTCAACCTCACAGCAGCAGCGTTCATCCTGGCCTTCATCAGCCTGACCGGT  528
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCTACACTGTCAACCTCACAGCAGCAGCGTTCATCCTGGCCTTCATCAGCCTGACCGGT  480

Qy   529 ACCTGCGCCATTGTCCACGCCACCAACTCTCGGGCCAAGGAAACTGCTGGCTGAAAAGG 588
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACCTGCTATTGTCCACGCCACCAACTCTCGGGCCAAGGAAACTGCTGGCTGAAAAGG    540

Qy   589 CAGTTCATGTGGCGTTCGGATGTGGTGTTCAATTCAGCATATACCTGACTTCATCTTTG 648
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CAGTTCATGTGGCGTTCGGATGTGGTGTTCAATTCAGCATATACCTGACTTCATCTTTG 600

Qy   649 CCGACGTCAGCCAGGCGGGACATCAGTCAGGGGGGGATGGGGGGGATGACAGCAGTACATC 708
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CCGACGTCAGCCAGGCGGGACATCAGTCAGGGGGGGATGGGGGGGATGACAGCAGTACATC 660

Qy   709 ACCCCGATAGCCGTGATGGTGGTGTTCAATTCCAGCAGATATAATGGTGGGTCTTCATCC 768
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCCGATAGCCGTGATGGTGGTGTTCAATTCCAGCAGATATAATGGTGGGTCTTCGTCC 720

Qy   769 TGCCCGGCATCGTCATCCTCTCCGTTACTGCATCATCATCCTCTAAGCTGTCACACTCCA 828
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TGCCCGGCATCGTCATCCTCTCCGTTACTGCATCATCATCCTCTAAGCTGTCACACTCCA 780

Qy   829 AGGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGTACATCCTCATCCTCAGCTTCTTG 888
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGTACATCCTCATCCTCAGCTTCTTG 840

Qy   889 CCTGCTGGCTGCCATATTATGTGGGGATCAGCATGCATGCATCCTCCTTCATCCTTTGGGAGTCA 948
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CCTGCTGGCTGCCATATTATGTGGGGATCAGCATGCATGCATCCTCCTTCATCCTTTGGGGGTCA 900

Qy   949 TCAAGCAAGGATGTGACTTGCACAAGTGGACATCCCATCACAGAGGCCC           1008
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCAAGCAAGGATGTGACTTGCACAAGTGGACATCCCATCACAGAGGCCC            960

Qy  1009 TCGCCCTTCTTCCACTGTGCCTGAACCCATCCTCTCATGCCGTTCCTCGGGGCCAAGTTCA 1068
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TCGCCCTTCTTCCACTGTGCCTGAACCCATCCTCTCATGCCGTTCCTCGGGGCCAAGTTCA 1020

Qy  1069 AAAGCTCTGCCCAGCATGCACTCAACTCCATGAGCAGAGAGGCTCCAGCCTCAAGATCCTTT 1128
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 AAAGCTCTGCCCAGCATGCACTCAACTCCATGAGCAGCAGAGAGGCTCCAGCCTCAAGATCCTTT 1080

Qy  1129 CCAAAGGAAAGGGGGGGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTTC  1188
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CCAAAGGAAAGGCGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTTC 1140

Qy  1189 ACTCCAGCTAACCCTTATGCAAAGACTTATATA---ATATATATATATATGATAAAG    1244
         ||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db  1141 ACTCCAGCTAACACTTATGCAAAGACATATAATAATATATATATATATATGATAAAG     1200

Qy  1245 AACTTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCCAGTCTGTACAGTTTT   1304
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 AACTTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCCAGTCTGTACAGTTTT   1260

Qy  1305 TTTTTTTTTTAATTGACTGTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTTGACTT 1364
         ||||||||         |||||||||||||||||||||||||||||||||||||||||
Db  1261 TTTTTT-----ATTGACTGTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTTGACTT 1315

Qy  1365 AATTTATATAATA----TTGTTTTTTGTTTGTTCATGTGAATGAGGCGTCTAGGCAGGA 1420
         ||||||||||||     |||||||||||||||||||||||||||||||||||||||||
Db  1316 AATTTATATAATACCTTTTTTTTGTTTGTTCATGTGAATGAGTGTCTAGGCACGAGA   1375

Qy  1421 CCTGTGGCCAAGTTCTTAGTAGCTTAATCTGTTGTGTCTGAGACTGTGAGAGGA      1480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1376 CCTGTGGCCAAGTTCTTAGTAGCTTAATCTGTTGTGTCTGAGACTGTGAGAGGA       1435

Qy  1481 AGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCT 1540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1436 AGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCT 1495

Qy  1541 GTTGCTGCATAATCCTTCATTCCGAGGAGCACCCCACCCCACCCCACCCCA          1600
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1496 GTTGCTGCATAATCTTCATTCCGAGGAGCACCCCACCCCACCCCACCCCA           1555

Qy  1601 TTCTTAAATTGTTGGTTATGCGTGTGATGGTTGTTTTTTTTT---GTTGTG          1657
         |||||||||||||||||||||||||||||||||||||         |||||
Db  1556 TTTTTTAAATTGTTGGTTATGCGTGTGATGGTTGTTTTTTTTTTTTTTGTTTTT       1615

Qy  1658 TTGTTGTTTTTTTTTCTGTAAAAGATGGCACTTAAAACCAAATGGTGGTAG          1717
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1616 TGTTTTTGTTTTTTTTTTCTGTAAAAGATGGCACTTAAAACCAAATGGTGGTAG        1675

Qy  1718 AAATGCTGGGGTTTTTTTGTTTGTTGTTTGTTTCAGTTTTCAAGAGTGACTTCA       1777
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1676 AAATGCTGGGGGTTTTTTTGTTTGTTGTTTGTTTCAGTTTTCAAGAGTGACTTCA      1735

Qy  1778 GTCCCTACAAATGTACAGTCTTGTATTACATTGTAATAAAGTCAATGATAAAACTTAAA  1837
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1736 CTCCCTACAAATGTACAGTCTTGTATTACATTGTAATAAAGTCAATGATAAAACTTAAA  1795

Qy  1838 AAATGCTGGGGTTTTTTTTGTTTTTTGTTTGTTGTTTGTTT  AAATGCTGGGGT      1843
         ||||||
Db  1796 AAATGCTGGGGT                                                1801
```

RESULT 4

```
LOCUS       AC145861              175370 bp    DNA     linear   HTG 01-AUG-2003
DEFINITION  Pan troglodytes chromosome UNK clone RP43-21D9, *** SEQUENCING IN
            PROGRESS ***, 39 unordered pieces.
ACCESSION   AC145861
VERSION     AC145861.1  GI:33386791
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 175370)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 175370)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT     ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information --------------
            Center project name: C_PT021D9
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 39 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1:     5345: contig of 5345 bp in length
            *     5346:     5445: gap of unknown length
            *     5446:    17195: contig of 11750 bp in length
            *    17196:    17295: gap of unknown length
            *    17296:    33250: contig of 15955 bp in length
            *    33251:    33350: gap of unknown length
            *    33351:    45646: contig of 12296 bp in length
            *    45647:    45746: gap of unknown length
            *    45747:    61992: contig of 16246 bp in length
            *    61993:    62092: gap of unknown length
            *    62093:    63423: contig of 1331 bp in length
            *    63424:    63523: gap of unknown length
            *    63524:    64666: contig of 1143 bp in length
            *    64667:    64766: gap of unknown length
            *    64767:    65994: contig of 1228 bp in length
            *    65995:    66094: gap of unknown length
            *    66095:    67430: contig of 1336 bp in length
            *    67431:    67530: gap of unknown length
            *    67531:    68982: contig of 1452 bp in length
            *    68983:    69082: gap of unknown length
            *    69083:    70891: contig of 1809 bp in length
            *    70892:    70991: gap of unknown length
            *    70992:    72172: contig of 1181 bp in length
            *    72173:    72272: gap of unknown length
            *    72273:    73614: contig of 1342 bp in length
            *    73615:    73714: gap of unknown length
            *    73715:    75185: contig of 1471 bp in length
            *    75186:    75285: gap of unknown length
            *    75286:    76854: contig of 1569 bp in length
            *    76855:    76954: gap of unknown length
            *    76955:    78413: contig of 1459 bp in length
            *    78414:    78513: gap of unknown length
            *    78514:    79964: contig of 1451 bp in length
            *    79965:    80064: gap of unknown length
            *    80065:    81745: contig of 1681 bp in length
            *    81746:    81845: gap of unknown length
            *    81846:    83417: contig of 1572 bp in length
            *    83418:    83517: gap of unknown length
            *    83518:    84910: contig of 1393 bp in length
            *    84911:    85010: gap of unknown length
            *    85011:    86788: contig of 1778 bp in length
            *    86789:    86888: gap of unknown length
            *    86889:    88992: contig of 2004 bp in length
            *    88893:    88992: gap of unknown length
            *    88993:    90670: contig of 1678 bp in length
            *    90671:    90770: gap of unknown length
            *    90771:    93207: contig of 2437 bp in length
            *    93208:    93307: gap of unknown length
            *    93308:    96658: contig of 3351 bp in length
            *    96659:    96758: gap of unknown length
            *    96759:   100576: contig of 3818 bp in length
            *   100577:   100676: gap of unknown length
            *   100677:   104988: contig of 4312 bp in length
            *   104989:   105088: gap of unknown length
            *   105089:   108893: contig of 3805 bp in length
            *   108894:   108993: gap of unknown length
            *   108994:   112632: contig of 3639 bp in length
            *   112633:   112732: gap of unknown length
            *   112733:   116338: contig of 3596 bp in length
            *   116329:   116428: gap of unknown length
            *   116429:   120980: contig of 4552 bp in length
            *   120981:   121080: gap of unknown length
            *   121081:   126821: contig of 5741 bp in length
            *   126822:   126921: gap of unknown length
            *   126922:   131331: contig of 4410 bp in length
            *   131332:   131431: gap of unknown length
            *   131432:   138705: contig of 7274 bp in length
            *   138706:   138805: gap of unknown length
            *   138806:   145169: contig of 6364 bp in length
            *   145170:   145269: gap of unknown length
            *   145270:   150696: contig of 5427 bp in length
            *   150697:   150796: gap of unknown length
            *   150797:   159988: contig of 9192 bp in length
            *   159989:   160088: gap of unknown length
            *   160089:   167244: contig of 7156 bp in length
            *   167245:   167344: gap of unknown length
            *   167345:   175370: contig of 8026 bp in length.
FEATURES             Location/Qualifiers
     source          1..175370
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="UNK"
                     /clone="RP43-21D9"
     misc_feature    1..5345
                     /note="assembly_name:Contig100"
     misc_feature    5446..17195
                     /note="assembly_name:Contig101"
     misc_feature    17296..33250
                     /note="assembly_name:Contig102"
     misc_feature    33351..45646
                     /note="assembly_name:Contig103"
     misc_feature    45747..61992
                     /note="assembly_name:Contig104"
     misc_feature    62093..63423
                     /note="assembly_name:Contig16"
     misc_feature    63524..64666
                     /note="assembly_name:Contig3"
     misc_feature    64767..65994
                     /note="assembly_name:Contig45"
     misc_feature    66095..67430
                     /note="assembly_name:Contig48"
     misc_feature    67531..68982
                     /note="assembly_name:Contig5"
     misc_feature    69083..70891
                     /note="assembly_name:Contig56"
     misc_feature    70992..72172
                     /note="assembly_name:Contig60"
     misc_feature    72273..73614
```

```
Db  168142  TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC  168201
                    ||||||||||||||||||||| ||| |||| |||||||||||||||||||||||||||||
Qy     361  ACCTGTCAGTGCTGCGACCTCCTTTGTCATCACACTCCTTTCTGGGCAGTTGATGCCA     420
                    ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  168202  ACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCTTTCTCGGCAGTTGATGCCA  168261
Qy     421  TGGCTGACTGGTACTTTGGGAAATTTTGTGTTAAGGCTGTCCATATCATCTACACTGTCA     480
                    |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  168262  TGGCTGACTGGTACTTTGGGAAATTTTGTGTTAAGGCTGTCCATATCATCTACACTGTCA  168321
Qy     481  ACCTCACAGCAGCGTTCTCATCCTGGCCTCATCAGCGTGACCGGTACCTCGCCATTG     540
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168322  ACCTCACAGCAGCGTTCTCATCCTGGCCTCATCAGCCTGACCGGTACCTCGCCATTG  168381
Qy     541  TCCACGCGCCACCAACAGTCAAAGCCAAGGAAACTGCTGGCTGGAAAAGGCAGTCTATGTGG     600
                    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168382  TCCACGCGCCACCAACAGTCAGAGGCCAAGGAAACTGCTGGCTGGAAAAGGCAGTCTATGTGG  168441
Qy     601  GGGTCTGGATCCCAGCCCTCCTGACTACTACCTGACTTCATCTTTGCCGACGTCAGCC     660
                    |||||| |||||||||||||| |||||||||||||||||||| |||||||||||||||||
Db  168442  GCGTCTGGATCCCAGCCCTCCTGACTACTACCTGACTTCATCTTTGCCGACGTCAGCC  168501
Qy     661  AGGGGGACATCAGTCAGGGGGATGACAGGTACATCAATTCAGCATATAAATGGTGGGTC     720
                    |||| ||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  168502  AGGGGGACATCAGTCAGGGGGATGACAGGTACATCAGTTCAGCATATAAATGGTGGGTC  168561
Qy     721  TGTGGATGGTGGTGTTCAATTCCAGCATCATCTCTAAGGTGTCACACTCGTCAAGGGCCACCAGA     780
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168562  TGTGGATGGTGGTGTTCAATTCCAGCATCATCTCTAAGGTGTCACACTCGTCAAGGGCCACCAGA  168621
Qy     781  TCATCCTCTCGTTACTGCATCATCATCTCTAAGCTGTCATCCTCCAGTCTTCTTTGCCTCTGGCTGC     840
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168622  TCATCCTCTCGTTACTGCATCATCATCTCTAAGCTGTCATCCTCCAGTCTTCTTTGCCTCTGGCTGC  168681
Qy     841  AGCGCAAGGCCCCTCAAGACGACAGTCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGAT     900
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168682  AGCGCAAGGCCCCTCAAGACGACAGTCATCGACTCCTTCATCCTTTGGGGTCATCAAGCAAGGAT  168741
Qy     901  CATATTATGTGGGATCAGCATCGACTCGACTCGCTTCATCCTTTCATCCTTTGGGGGTCATCAAGCAAGGAT     960
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168742  CATATTATGTGGGATCAGCATCGACTCGACTCGCTTCATCCTTTCATCCTTTGGGGGTCATCAAGCAAGGAT  168801
Qy     961  GTGACTTCGAGAGCATTGTGCACAAGTGGATCCTCATCACAGAGAGGCCCTGCCTTCTTCTCC     1020
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168802  GTGACTTCGAGAGCATCGTGCACAAGTGGATCCTCATCACAGAGAGGCCCTGCCTTCTTTCC  168861
Qy    1021  ACTGTTGCCTGAACCCATCCTCTATGCCTTCTCGGGGCCAAGTCAAAAGCTCTCGCCC     1080
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168862  ACTGTTGCCTGAACCCATCCTCTATGCCTTCTCGGGGGCCAAGTCAAAAGCTCTCGCCC  168921
Qy    1081  AGCATGCACTCAACTCACTCCAGCCTCCAGCGCAGAGGCTCCAGCCTCCAGATCCTTCCAAAAGGAAAGC     1140
                    |||||||||||||| ||||||||||| ||||||||||||||||||||||||||||||||
Db  168922  AGCATGCACTCAACTCACTCCAGCGACAGAGGCTCCAGCCTCCAGATCCTTCCAAAAGGAAAGC  168981
Qy    1141  GGGGTGGACACTCTTCCGTCTCCACCGAGTCGAGATCCTCCAGTTTTCACTCCAGCTAAC     1200
                    |||||| |||| || || ||| |||| |||||||||||||||||||||||||||||||||
Db  168982  GGGGTGGACACTTCTTCCGTCTCCACCGAGTCGAGAATCCTCCAGTTTTCACTCCAGCTAAC  169041
Qy    1201  CCTTATGCAAAGACTTATATA---ATATATATATATATATATATGATAAAGAACTTTTTATG     1256
                    ||||||||||||| |||||||   |||||||||||||||||||||||||||||||||||||
Db  169042  ACTTATGCAAAGACATATATATATATATATATATATATATATATGATAAAGAACTTTTTATG  169101
Qy    1257  TTACACATTTCCAGATATAAGAGACTTGTACAGTTTTTTTTTTTTTTTAA     1316
                    |||||||||||||||||||||||||||||||||||| |||||      |||
Db  169102  TTACACATTTCCAGATATAAGAGACTTGTACAGTTTTTTTTT---A  169157
Qy    1317  TTGACTGTTGGGAGTTTATGTTCCTCCAGTTTTTGTGAGGTTTGACTTAATTTATATAAA     1376
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  169158  TTGACTGTTGGGAGTTTATGTTCCTCCAGTTTTTGTGAGGTTTGACTTAATTTATATAAA  169217
Qy    1377  TA----TTGTTTTTTGTTGTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGGCCAAG     1432
                              |||||||||||||||||||||||||||||||||||||||||||||||
Db  169218  TACTTTTTTTGTTGTTGTTGTTCATGTGAATGAGTGTCTAGGCAGGACCTGTGGCCAAG  169277
```

```
misc_feature  73715..75185    /note="assembly_name:Contig67"
misc_feature  75286..76854    /note="assembly_name:Contig68"
misc_feature  76955..78413    /note="assembly_name:Contig69"
misc_feature  78514..79964    /note="assembly_name:Contig73"
misc_feature  80065..81745    /note="assembly_name:Contig75"
misc_feature  81846..83417    /note="assembly_name:Contig77"
misc_feature  83518..84910    /note="assembly_name:Contig79"
misc_feature  85011..86788    /note="assembly_name:Contig80"
misc_feature  86889..88892    /note="assembly_name:Contig81"
misc_feature  88993..90670    /note="assembly_name:Contig82"
misc_feature  90771..93207    /note="assembly_name:Contig83"
misc_feature  93308..96658    /note="assembly_name:Contig84"
misc_feature  96759..100576   /note="assembly_name:Contig85"
misc_feature  100677..104988  /note="assembly_name:Contig86"
misc_feature  105089..108893  /note="assembly_name:Contig87"
misc_feature  108994..112632  /note="assembly_name:Contig88"
misc_feature  112733..116328  /note="assembly_name:Contig89"
misc_feature  116429..120980  /note="assembly_name:Contig90"
misc_feature  121081..126821  /note="assembly_name:Contig91"
misc_feature  126922..131331  /note="assembly_name:Contig92"
misc_feature  131432..138705  /note="assembly_name:Contig93"
misc_feature  138806..145169  /note="assembly_name:Contig94"
misc_feature  145270..150696  /note="assembly_name:Contig95"
misc_feature  150797..159988  /note="assembly_name:Contig96"
misc_feature  160089..167244  /note="assembly_name:Contig97"
misc_feature  167345..175370  /note="assembly_name:Contig98"
                               /note="assembly_name:Contig99"

ORIGIN

Query Match      85.4%;     Score 1602.6;   DB 2;   Length 175370;
Best Local Similarity  97.2%;  Pred. No. 0;
Matches 1678;  Conservative 0;  Mismatches 34;  Indels 15;  Gaps 4;
```

```
Qy     121  TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG  180
                    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  167962  TGTCCCGTGTTCCTTGCAGATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTTCTG  168021
Qy     181  GAGACTATGACTCCAACAGGAACCCTGCTTCCGGGATGAAAAACGTCCATTTCAATAGGA  240
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168022  GAGACTATGACTCCAACAGGAACCCTGCTTCCGGGATGAAAAACGTCCATTTCAATAGGA  168081
Qy     241  TCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCATGGCATTGG  300
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  168082  TCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCATGGCATTGG  168141
Qy     301  TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC  360
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy      1433 TTCTTAGTAGCTGTTTATCTGTGTGTGTAGGACTGTGAGAACTGTGAGAGGAAGAAACTGAACA 1492
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169278 TTCTTAGTAGCTGTTTATCTGTGTGTGTAGGACTGTGAGAACTGTGAGAGGAAGAAACTGAACA 169337

Qy      1493 TTCCAGAATGTGTGGGTAAATTGAATAAAGTTAGCCGTGATCCTCAGCTGTTGCTGCATAA 1552
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169338 TTCCAGAATGTGTGGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCATAA 169397

Qy      1553 TCTCTTCATTCCGAGGAGCACCCACCCCACCCCCACCCCCATTCTTAAATTGT 1612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169398 TCTCTTCATTCCGAGGAGCACCCACCCCACCCCCACCCCCATTCTTAAATTGT 169457

Qy      1613 TTGGTTATGCTGTGTGATGGTTTGTTTGGTTTTTTTT---GTTGTTGTTGTTTGTTTTTT 1669
             |||||||||||||||||||||||||||||        |||||||||||||| ||||||
Db    169458 TTGGTTATGCTGTGTGATGGTTTGTTTGGTTTTTTTTTTTTGTTGTTGTTTGTTTTTT 169517

Qy      1670 TTTTCTGTAAAAGATGGCACTTAAAAACCAAAGCCTGAAAATGTGGTAGAAATGCTGGGGT 1729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169518 TTTTCTGTAAAAGATGGCACTTAAAAACCAAAGCCTGAAAATGTGGTAGAAATGCTGGGGG 169577

Qy      1730 TTTTTTTGTTGTTGTTTGTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAAT 1789
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169578 TTTTTTTGTTGTTGTTTGTTTTTCAGTTTTCAAGAGTAGATTGATTGATTTCACTCCCTACAAAT 169637

Qy      1790 GTACAGTCTTGTATTACATTGTTAATAAAAGTCAATGATAAACTTAA 1836
             |||||||||||||||||||||||||||||||||||||||||||||||
Db    169638 GTACAGTCTTGTATTACATTGTTAATAAAAGTCAATGATAAACTTAA 169684
```

RESULT 5
AC147556
LOCUS   AC147556   185215 bp   DNA   linear   HTG 17-DEC-2003
DEFINITION Mus musculus chromosome UNK clone RP23-480N14, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC147556
VERSION   AC147556.1 GI:39979584
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 185215)
AUTHORS   Wilson,R.K.
TITLE   The sequence of Mus musculus clone
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 185215)
AUTHORS   Wilson,R.K.
TITLE   Direct Submission
JOURNAL   Submitted (17-DEC-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT   ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            ----------- Project Information -----------
            Center project name: M_BA0480N14
            --------------- Summary Statistics ---------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 194611 bases at least Q40
            Consensus quality: 196848 bases at least Q30
            Consensus quality: 198294 bases at least Q20
            Insert size: 252319; agarose-fp
            Insert size: 252319; sum-of-contigs
            Quality coverage: 13.47 in Q20 bases; agarose-fp
            Quality coverage: 9.68 in Q20 bases; sum-of-contigs

```
* ----------------------------------------------------------------
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* ----------------------------------------------------------------
*        1        1119: contig of 1119 bp in length
*     1120        1219: gap of unknown length
*     1220        2637: contig of 1418 bp in length
*     2638        2737: gap of unknown length
*     2738        5059: contig of 2322 bp in length
*     5060        5159: gap of unknown length
*     5160        6203: contig of 1044 bp in length
*     6204        6303: gap of unknown length
*     6304       11814: contig of 5511 bp in length
*    11815       11914: gap of unknown length
*    11915       18453: contig of 6539 bp in length
*    18454       18553: gap of unknown length
*    18554       33684: contig of 15131 bp in length
*    33685       33784: gap of unknown length
*    33785       57713: contig of 23929 bp in length
*    57714       57813: gap of unknown length
*    57814       84551: contig of 26738 bp in length
*    84552       84651: gap of unknown length
*    84652      122285: contig of 37634 bp in length
*   122286      122385: gap of unknown length
*   122386      185215: contig of 62830 bp in length.
```

FEATURES   Location/Qualifiers
     source   1..185215
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP23-480N14"
     misc_feature   1..1119
                     /note="assembly_name:Contig51"
     misc_feature   1220..2637
                     /note="assembly_name:Contig103"
     misc_feature   2738..5059
                     /note="assembly_name:Contig122"
     misc_feature   5160..6203
                     /note="assembly_name:Contig124"
     misc_feature   6304..11814
                     /note="assembly_name:Contig126"
     misc_feature   11915..18453
                     /note="assembly_name:Contig127"
     misc_feature   18554..33684
                     /note="assembly_name:Contig128"
     misc_feature   33785..57713
                     /note="assembly_name:Contig129"
     misc_feature   57814..84551
                     /note="assembly_name:Contig130"
     misc_feature   84652..122285
                     /note="assembly_name:Contig131"
     misc_feature   122386..185215
                     /note="assembly_name:Contig132"
ORIGIN

```
Query Match        85.4%;  Score 1602.6;  DB 2;  Length 185215;
Best Local Similarity  97.2%;  Pred. No. 0;
Matches 1678;  Conservative  0;  Mismatches  34;  Indels  15;  Gaps  4;

Qy   121 TGGAACCGATCAGTGTGAGTATATACACTTCGATAACTACTTCGAAGAAGTGGGGTCTG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91301 TGTCCCGTTCCTTGCAGATATACACTTCTGATAACTACTTCGAAGAAGTGGGGTTCTG 91360

Qy   181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91361 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGA 91420
```

```
  241 TCTTCCTGCCCACCACTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 300
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91421 TCTTCCTGCCCACCAATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 91480

  301 TGATCCTGGTCATGGGTTACCAGAGAAGCTAAGGACATGACGGACAAGTACCGGCTGC 360
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91481 TGATCCTGGTCATGGGTTACCAGAGAAGCTAAGGACATGACGGACAAGTACCGGCTGC 91540

  361 ACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCTTCTGGGCAGTTGATGCCA 420
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91541 ACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCTTCTGGGCAGTTGATGCCA 91600

  421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91601 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 91660

  481 ACCTCACACAGCGGTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTG 540
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91661 ACCTCACACAGCGGTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCTATTG 91720

  541 TCCACCCACCAACAGTCAAAGCCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTCG 600
      ||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||
91721 TCCACCGCCACCAACAGTCAGGTCAGCCCTCCTCGACTATACCTGCTGGCTGAAAAGGCAGTCTATGTCG 91780

  601 GCGTCTGGATCCCAGCCCTCCTGACTATACCTGCTGGCTGAAAAGGCAGTCTATGTGCCGACGTCAGCC 660
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91781 GCGTCTCGGATCCCAGCCCTCCTGACTATACCTGCTGGCTGAAAAGGCAGTCTATGTGCCGACGTCAGCC 91840

  661 AGGGGGACACATCAGTCAGGGGGGATGACAGGTACACATCGTGACCGGCCTTTACCCGTTACCCCGATAGCC 720
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91841 AGGGGGACACATCAGTCAGGGGGGATGACAGGTACACATCGTGACCGGCCTTTACCCGTTACCCCGATAGCC 91900

  721 TGTGGATGGTGGTTTCAATTCAGCATATAATGTGGGCTCTCATCCTCGTCCTGCCCGGCATCG 780
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91901 TGTGGATGGTGTGTTTCAATTCAGCATATAATGTGGGCTCTCATCCTCGTCCTGCCCGGCATCG 91960

  781 TCATCCTCTCTGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCACCAGA 840
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
91961 TCATCCTCTCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGA 92020

  841 AGCGCAAGGCCCCAGAAGACACAGCAGTCATCATCCTAGCTTCTTGCCTGCTGGCTGC 900
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92021 AGCGCAAGGCCCCAAGAAGACACAGCAGTCATCATCCTAGCTTCTTGCCTGCTGGCTGC 92080

  901 CATATTATGTGGGATCAGCATGACATCGACTCCTTTTGGGAGTCATCAAGCAAGGAT 960
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92081 CATATTATGTGGGATCAGCATCAGCATGACATCGACTCCTTTTGGGGGTCATCAAGCAAGGAT 92140

  961 GTGACTTCGAGAGCATTGTGCACAAGTGAGATCCATCACAGAGGCCCTCGCCTTCTTCC 1020
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92141 GTGACTTCGAGAGCATCGTGCACAAGTGGATCCATCACAGAGGCCCTCGCCTTCTTCC 92200

 1021 ACTGTTGCCTGAACCCATCCTCTATGCCTTCCTCGGGGGCCAAGTTCAAAAGCTCTGCCC 1080
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92201 ACTGTTGCCTGAACCCATCCTCTATGCCTTCCTCGGGGGCCAAGTTCAAAAGCTCTGCCC 92260

 1081 AGCATGCACTCAACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGC 1140
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92261 AGCATGCACTCAACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGC 92320

 1141 GGGGTGACACTCCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAAC 1200
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92321 GGGGTGGACACTCCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAAC 92380

 1201 CCTTATGCAAAGACTTATATA---ATATATATATATATGATAAAGAACTTTTTTATG 1256
      ||||| |||| |||||||||     |||||||||||||||||||||||||||||||||
92381 ACTTATGCAAAGACATATATATATAATATAATATATATATATGATAAAGAACTTTTTTATG 92440

 1257 TTACACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTAA 1316
      ||||| ||||||||||||||||||| |||||||||||||||||||||||||||||
92441 TTACACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTTTTTTT---A 92496

 1317 TTGACTGTTGGGAGTTTATGTTCCTCCTCAGTTTTTGTGAGGTTTGACTTAATTTATATAAA 1376
```

```
92497 TTGACTGTTGGGAGTTTATGTTCCTCCTCAGTTTTTGTGAGGTTTGACTTAATTTATATAAA 92556

 1377 TA---TTGTTTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGTCGGCCAAG 1432
      ||   ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
92557 TACTTTTTTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGTCGGCCAAG 92616

 1433 TTCTTAGTAGCTGTTTATCTGTGTGAGGACTGTAGGACTGTAGGAGGAGAAACTGAACA 1492
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92617 TTCTTAGTAGCTGTTTATCTGTGTGAGGACTGTAGGACTGTAGGAGGAGAAACTGAACA 92676

 1493 TTCCAGAATGTGTGTAAATTGAATAAAGCTAGCCGGTGATCCTCAGCTGTTGCTGCATAA 1552
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92677 TTCCAGAATGTGTGTAAATTGAATAAAGCTAGCCGGTGATCCTCAGCTGTTGCTGCATAA 92736

 1553 TCTCTTCATTCCGAGGAGCACCCACCCCACCCCACCCCACCCCATTCTAAAATTGT 1612
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92737 TCTCTTCATTCCGAGGAGCACCCACCCCACCCCACCCCACCCCATTCTTAAAATTGT 92796

 1613 TTGGTTATGCTGTGTGGATGGTTTGTTGGTTTTTTT---GTTGTTGTTGTTTTT 1669
      ||||||||||||||||||||||||||||||| |||||    ||||||||||||||||||
92797 TTGGTTATGCTGTGTGGATGGTTTTGTTGGTTTTTTTTTTGTTTTTTGTTTTTTGTTTTT 92856

 1670 TTTTCGTAAAAGATGGCACTTAAAACCAAAGCCTGAAAATGGTGTAGAAATGGTGGGGT 1729
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
92857 TTTTCGTAAAAGATGGCACTTAAAACCAAAGCCTGAAAATGGTGTAGAAATGGTGGGGG 92916

 1730 TTTTTTTGTTGTTGTTGTTTTTCAGTTTTCAAGAGTAGATTGATTTCACTCCCTACAAAT 1789
      |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
92917 TTTTTTTTGTTGTTGTTTTTCAGTTTTCAAGAGTAGATTGATTTCACTCCCTACAAAT 92976

 1790 GTACAGTCTTGTATTACATTGTTAATAAAGTCAATGATGATAAAACTTAA 1836
      |||||||||||||||||||||||||||||||||||||||||||||||||||
92977 GTACAGTCTTGTATTACATTGTTAATAAAGTCAATGATGATAAAACTTAA 93023
```

RESULT 6
MMLESTRPT
LOCUS   MMLESTRPT   1575 bp   mRNA   linear   ROD 11-NOV-1996
DEFINITION   M.musculus mRNA for leukocyte-derived seven transmembrane domain
               receptor.
ACCESSION   X99582
VERSION   X99582.1 GI:1666648
KEYWORDS   lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Moepps,B., Frodl,R., Kessler,H. and Gierschik,P.
  TITLE   cDNA cloning and genomic organization of a leukocyte-derived seven
               transmembrane domain receptor (LESTR) from mouse: a murine
               homologue of the human HIV-1 entry cofactor fusin
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1575)
  AUTHORS   Moepps,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,
               Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK   Revised by author 11-NOV-96
COMMENT   On Nov 12, 1996 this sequence version replaced gi:1657351.
FEATURES   Location/Qualifiers
     source   1..1575
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="B6CBA"
                     /db_xref="taxon:10090"
                     /haplotype="diploid"
                     /sex="female"
                     /tissue_type="thymus"
                     /dev_stage="6-8 weeks old"
     gene   1..1575
                     /gene="lestr"

```
CDS          90...1169
             /gene="lestr"
             /codon_start=1
             /product="leukocyte-derived seven transmembrane domain
             receptor"
             /protein_id="CAA67894.1"
             /db_xref="GI:1666649"
             /db_xref="GOA:P70658"
             /db_xref="SWISS-PROT:P70658"
             /translation="MEPISVSIYTSDNYSEVGSGYDSNKEPCFRDENVHFNRIFLP
             TIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
             DWYFGKFLCKAVHIIYTVNLYSSVLLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
             GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFOHIMVGLVLP
             GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
             IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFIGAKFKSSAQHALNSMSRGSSLK
             ILSKGKRGHSSVSTSESBSSSFHSS"
polyA_signal 1498...1503
             /gene="lestr"
ORIGIN
```

Query Match   78.6%; Score 1474.8; DB 10; Length 1575;
Best Local Similarity 98.4%; Pred. No. 5.1e-296;
Matches 1525; Conservative 0; Mismatches 12; Indels 13; Gaps 3;

```
Qy    50  GCAGTAGCAGTGACCCTGAGCGTTGGTGTCTCCGTAACCACCACGGCTGTAGAGC  109
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    20  GCAGTAGCAGTGACCCCTGAGCGTTGGTGTCTCCGTAACCACCACGGCTGTAGAGC  79

Qy   110  GAGTGTTGCCATGGAACCGATCAGTGTGAGTAGTATACACTTCTGATAACTACTCTGAAGA  169
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    80  GAGTGTTGCCATGGAACCGATCAGTGTGAGTAGTATACACTTCTGATAACTACTCTGAAGA  139

Qy   170  AGTGGGGTCGGAGACTTCCAACAAGGAACCCGCTTCCGGGATGAAGAAACGTCCA  229
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   140  AGTGGGGTTCGGAGACTTCCAACAAGGAACCCTGCTTCCGGGATGAAGAAACGTCCA  199

Qy   230  TTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCATCATGGCATAGTCGG  289
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   200  TTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCATCATGGCATAGTCGG  259

Qy   290  CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGAAGTCATGAGCGTA  349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   260  CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGAAGTAAGGACGACGAA  319

Qy   350  GTACCGGCTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCCCTTCTGGGC  409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   320  GTACCGGCTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCCCTTCTGGGC  379

Qy   410  AGTTGATGCCATGGCCAGCTGACTGGTACTTTGGGAAATTTGTGTAAGGCTGTCCATATCAT  469
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   380  AGTTGATGCCATGGCCAGCTGACTGGTACTTTGGGAAATTTGTGTAAGGCTGTCCATATCAT  439

Qy   470  CTACACTGTCAACCTACAGCACGCCACCACTCTACAGCGCGTTCATCAGCCTGGACCCGTA  529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   440  CTACACTGTCAACCTACAGCACGCCACCACTCTACAGCGCGTTCATCAGCCTGGACCCGTA  499

Qy   530  CCTCGCCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGAAAACTGCTGGCTGAAAAGGC  589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   500  CCTCGCTATTGTCCACGCCACCAACAGTCAGAGGCCAAGGAAAACTGCTGGCTGAAAAGGC  559

Qy   590  AGTCTATGTGGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTTCATCTTTGC  649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   560  AGTCTATGTGGGGCGTCTGGATCCCAGCCCTCCTCGACTATACCTTCATCTTTGC  619

Qy   650  CGACGTCAGCCAGGGGGACACTCAGTCAGGGGATGACAGAGTACATCGTGACCGCCTTTA  709
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   620  CGACGTCAGCCAGGGGGACATCAGTCAGGGGATGACAGAGTACATCGTGACCGCCCTTTA  679

Qy   710  CCCGGATAGCCTGTGGATGGTGGTTCAATTCACCAGCATATAAATGTGGGTCATCCT  769
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   680  CCCGGATAGCCTGTGGATGGTGGTTTCAATTCCAGCATATAAATGTGGGTCCGTCCT  739

Qy   770  GCCCGGCATCGTCATCCTCCGTACTGCATCATCATCACTTCTAAGCTGTCACACTCCAA  829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   740  GCCCGGCATCGTCATCCTCCGTGTTACTGCATCATCATCACTTCTAAGCTGTCACACTCCAA  799

Qy   830  GGGCCACCAGAAGCGCAAGGCCAAGGCCCTCAAGACCAGTACAGTCATCCTCATCTTTCTTGC  889
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   800  GGGCCACCAGAAGCGCAAGGCCAAGGCCCTCAAGACGACAGTCATCCTCATCTTTCTTGC  859

Qy   890  CTGCTGCTGCCATATTATGTGGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCAT  949
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   860  CTGCTGCTGCCATATTATGTGGGGATCAGCATCGACTCCTTCATCCTTTTGGGGTCAT  919

Qy   950  CAAGCAAGGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGCCCT  1009
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   920  CAAGCAAGGATGTGACTTCGAGAGCATCGTGCACAAGTGGATCTCCATCACAGAGCCCT  979

Qy  1010  CGCCTTCTTCCACTGTGCCTGCCTGCGAACCCCATCCTCATGCCTCTATGCCTTCCTCGGGGCCAAGTTCAA  1069
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   980  CGCCTTCTTCCACTGTGCCTGCCTGCGAACCCCATCCTCATGCCTCTATGCCTTCCTCGGGGCCAAGTTCAA  1039

Qy  1070  AAGCTGCCCAGCATGCACTCAACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTTCA  1129
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1040  AAGCTGCCCAGCATGCACTCAACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTTC  1099

Qy  1130  CAAAGGAAAGGCGGGGGGTGGACACTCTTCCGTCCCACGGAGTCAGAATCCTCCAGTTTTCA  1189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1100  CAAAGGAAAGGCGGGGGGTGGACACTCTTCCGTCCCACGGAGTCAGAATCCTCCAGTTTTCA  1159

Qy  1190  CTCCAGTAACCCTTATGCAAAGACTTATATA---ATATATATATAATATATGATAAAGA  1245
          |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db  1160  CTCCAGTAACACTTATGCAAAGACATTATATATATAATATATATATAATATATGATAAAGA  1219

Qy  1246  ACTTTTTATGTTACACATTT-----TT  1301
          ||||||||||||||||||||     ||
Db  1220  ACTTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAGCTGTGTACAGTTTTTT  1279

Qy  1302  TTTTTTTTTTTTAATTGACTGTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGA  1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1280  TTTTTTTTTTTTTTATTGACTGTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGA  1339

Qy  1362  CTTAATTTATATAAATA-----TTGTTTTTTGTTTGTTCATGTGAATGAGCGTCTAGGC  1416
          |||||||||||||||||     |||||||||||||||||||||||||||||||||||||
Db  1340  CTTAATTTATATAAATACCTTTTTTGTTTTTTGTTTGTTCATGTGAATGAGTGTCTAGGC  1399

Qy  1417  AGGACCTGTGGCCAAGTTCTTAGTAGCTGTGTTTATCTGTGTGTGGTAGGACTGTGTAGAACTGTAG  1476
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1400  AGGACCTGTGGCCAAGTTCTTAGTAGCTGTGTTTATCTGTGTGTGGTAGGACTGTGTAGAACTGTAG  1459

Qy  1477  AGGAAGAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAAGCTAGGCCGTGATCCTC  1536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1460  AGGAAGAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAAGCTAGCCGTGATCCTC  1519

Qy  1537  AGCTGTTGCTGCATAATCCTTCATTCCGAGGAGCACCCCCACCC  1586
          ||||||||||||||||||||||||||||||||||||||||||||
Db  1520  AGCTGTTGCTGCATATATCTTCATTCCGAGGAGCACCCCCACCC  1569
```

RESULT 7
MMLESTRGN
LOCUS   MMLESTRGN   3770 bp   DNA   linear   ROD 02-APR-1997
DEFINITION M.musculus gene encoding leukocyte-dervied seven transmembrane
          domain receptor, strain B6.
ACCESSION   X99581
VERSION   X99581.1 GI:1924959
KEYWORDS   lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
  AUTHORS   Moepps,B., Frodl,R., Kessler,H. and Gierschik,P.
  TITLE   cDNA cloning and genomic organization of a leukocyte-derived seven
         transmembrane domain receptor (LESTR) from mouse: a murine
         homologue of the human HIV-1 entry cofactor fusin

```
JOURNAL    Unpublished
REFERENCE  2
AUTHORS    Moepps,B.
TITLE      Direct Submission
JOURNAL    Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,
           Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK     Revised by [3]
REFERENCE  3  (bases 1 to 3770)
AUTHORS    Moepps,B.
TITLE      Direct Submission
JOURNAL    Submitted (02-APR-1997) B. Moepps, Universitaet Ulm,
           Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
COMMENT    On Apr 3, 1997 this sequence version replaced gi:1666646.
FEATURES            Location/Qualifiers
     source         1..3770
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /strain="B6"
                    /db_xref="taxon:10090"
                    /haplotype="diploid"
                    /tissue_type="thymus"
                    /dev_stage="6-8 weeks old"
     exon           <1..110
                    /number=1
     gene           join(90..110,2306..3364)
                    /gene="lestr"
     CDS            join(90..110,2306..3364)
                    /gene="lestr"
                    /codon_start=1
                    /product="leukocyte-derived seven transmembrane domain
                    receptor"
                    /protein_id="CAA67893.1"
                    /db_xref="GI:1666647"
                    /db_xref="GOA:P70658"
                    /db_xref="SWISS-PROT:P70658"
                    /translation="MEPISVSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLP
                    TIYFIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                    DWYFGKFLCKAVHIIYTVNLYSSVLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                    GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQHIMVGLVLP
                    GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
                    IKQGDFPESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                    ILSKGKRGGHSSVSTESESSSFHSS"
     misc_feature   105..2305
                    /gene="lestr"
                    /note="splice variant/intron"
     intron         111..2305
                    /gene="lestr"
     exon           2306..>3770
                    /number=2
     polyA_signal   3693..3698
ORIGIN

Query Match           73.8%;   Score 1384.6;  DB 10;  Length 3770;
Best Local Similarity 97.5%;             Pred. No. 3.2e-277;
Matches 1442;  Conservative    0;  Mismatches 24;  Indels 13;  Gaps  3;

Qy  121 TGGAACCGATCAGTGTGAGTATACACTTCTGATAACTACTTCTGAAGAAGTGGGGTCTG  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2286 TGTCCCCTGTTTCCTTGCAGTCAACAAGATACACTTCTGATAACTACTTCTGAAGAAGTGGGGTCTG 2345

Qy  181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAACGTCCATTCAATAGGA  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2346 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAACGTCCATTCAATAGGA 2405

Qy  241 TCTTCCTGCCCACCATCTACTTCATCTTCCTGACTGGCATATCGGCAATGGATTGG  300
        |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||
Db 2406 TCTTCCTGCCCACCATCTACTTCATCATCCTGACTGGCATAGTCGGCAATGGATTGG 2465

Qy  301 TGATCCTGGTCATGGGCTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2466 TGATCCTGGTCATGGGCTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 2525

Qy  361 ACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCTTCTGGGCAGTTGATGCCA  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2526 ACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCTTCTGGGCAGTTGATGCCA 2585

Qy  421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCATATCATCATCACACTGTCA  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2586 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACTGTCA 2645

Qy  481 ACCTTACAGCAGCGTTCTCATCCTCGGCCTTCATCAGCTGGACCGGTACCTCGCCATTG  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2646 ACCTTACAGCAGCGTTCTCATCCTCGGCCTTCATCAGCTGGACCGGTACCTCGCCATTG 2705

Qy  541 TCCACGCCACCAACGTCAAAGGCCAAGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG  600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2706 TCCACGCCACCAACGTCAAAGGCCAAGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 2765

Qy  601 GCGTCTGGATCCAGCCCTCCTCGACTATACCTGACTTCATCTTGCCAGTGACGTCAGCC  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2766 GCGTCTGGATCCAGCCCTCCTCGACTATACCTGACTTCATCTTGCCAGTGACGTCAGCC 2825

Qy  661 AGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCGATAGCC  720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2826 AGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCGATAGCC 2885

Qy  721 TGTGGATGGTGTGTTCAATTCCAGCATATAATGGTGGGCTCATCCTCTGCCGGCATCG  780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2886 TGTGGATGGTGTGTTCAATTCCAGCATATAATGGTGGGCTCATCCTCTGCCGGCATCG 2945

Qy  781 TCATCCTCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTGTCACACTCAAGGGCCACCAGA  840
        || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2946 TCATCCTCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTGTCACACTCAAGGGCCACCAGA 3005

Qy  841 AGCGCAAGGCCCTCAAGACGACAGTGATCCTCATCCTAGCTTCTTTGCCTGCTGGCAAGGAT  900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3006 AGCGCAAGGCCCTCAAGACGACAGTGATCCTCATCCTAGCTTCTTTGCCTGCTGGCAAGGAT 3065

Qy  901 CATATATGTGGGATCAGCTGACTCTTTTGGGAGTCATCAAGCAAGGAT  960
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3066 CATATATGTGGGATCAGCTGACTCTTTTGGGGGTCATCAAGCAAGGAT 3125

Qy  961 GTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTGCCTTCTTCC 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3126 GTGACTTCGAGAGCATCGTGCACAAGTGGATCTCCATCACAGAGGCCCTGCCTTCTTCC 3185

Qy 1021 ACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTGGGGGCCAAGTCAAAGTCTCGCCC 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3186 ACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTGGGGGCCAAGTCAAAGTCTCGCCC 3245

Qy 1081 AGCACTGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAAGC 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3246 AGCACTGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAAGC 3305

Qy 1141 GGGGTGGACACTCTTCCGTCTCCGTCTCCAGTTTTCACTCCAGCTAAC  1200
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3306 GGGGTGGACACTCTTCCGTCTCCAGTTTTCACTCCAGCTAAC 3365

Qy 1201 CCTTATGCAAAGACTTATATA----ATATATATATATAGAGAGACTGGGTCTG 1256
        |||||||||||||||||||||    ||||||||||||||||||||||||||||
Db 3366 ACTTATGCAAAGACATATATATATATATATATAATATATAGAGAGACTTTTTATG 3425

Qy 1257 TTACACATTTCCAGATATAAGGAACGTCCATTCAATAGG----TTTTTTTTTT 1312
        ||||||||||||||||||||||||||||||||||||||    |||||||||||
Db 3426 TTACACATTTCCAGATATAAGGAGACGTCCATTCAATAGGTTTTTTTTTTTTTTT 3485

Qy 1313 TTAATTGACTGTTGGGAGTTTATGTTCCTCCAGTTTTTTGGGAGGTTGACTTAATTATA 1372
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3486 TTTATTGACTGTTGGGAGTTTATGTTCCTCCAGTTTTTTGGGAGGTTGACTTAATTATA 3545

Qy 1373 TAAATA----TTGTTTTTGTTTGTTTGTTTGTTTCATGAGAATGAGCGTCTAGGCAGGACCTGTGG 1427
        |||||    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3546 TAAATACCTTTTTTTTGTTTGTTTGTTTGTTTCATGAGAATGAGCGTCTAGGCAGGACCTGTGG 3605

Qy 1428 CCAAGTTCTTAGTAGCTGTTTATCGTGTGTGTAGGACTGTAGAACTGTAGAGAGGAAAACT 1487
```

```
Db   3606 CCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAAATTGAATAAAGTAGCCGTGATCCTCAGCTGTTGCTG 3665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1488 GAACATTCCAGAATGTGTGTAAATTGAATAAAGTAGCCGTGATCCTCAGCTGTTGCTG 1547
          |||||||||||||||||||||||||||||||
Db   3666 GAACATTCCAGAATGTGTGTAAATTGAATAAAGTAGCCGTGATCCTCAGCTGTTGCTG 3725
Qy   1548 CATAATCTCTTCATTCCGAGGAGCACCCCACCCCCACCC 1586
          |||||||||||||||||||||||||||||||||||||||
Db   3726 CATAATCTCTTCATTCCGAGGAGCACCCCACCCCCACCC 3764

RESULT 8
MMU59760
LOCUS       MMU59760          1223 bp    mRNA    linear    ROD 30-SEP-2002
DEFINITION  Mus musculus fusin homolog mRNA, complete cds.
ACCESSION   U59760
VERSION     U59760.1  GI:1527134
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1223)
AUTHORS     Heesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
TITLE       Cloning of the mouse fusin gene, homologue to a human HIV-1
            co-factor
JOURNAL     J. Immunol. 157 (12), 5455-5460 (1996)
MEDLINE     97113334
PUBMED      8955194
REFERENCE   2  (bases 1 to 1223)
AUTHORS     Heesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1996) Pathology, Harvard Medical School, 200
            Longwood Avenue, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..1223
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129/Sv"
                     /db_xref="taxon:10090"
     CDS             30..1109
                     /note="seven transmembrane spanning receptor;
                     chemokine-receptor like"
                     /codon_start=1
                     /product="fusin homolog"
                     /protein_id="AAB07725.1"
                     /db_xref="GI:1527135"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                     GIVILSYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTSESSSFHSS"
ORIGIN

Query Match    64.8%;  Score 1216.6;  DB 10;  Length 1223;
Best Local Similarity  99.7%;  Pred. No. 2.4e-242;
Matches 1219;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy     91 ACCACCACGGCTGTAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTT 150
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACCACGGCTGTAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTT 60

Qy    151 CTGATAACTACTCTGAAGAAGTGGGTCTGGAAGGATGACTCCAACAAGAACCCTGCT 210
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGATAACTACTCTGAAGAAGTGGGTCTGGAAGGATGACTCCAACAAGAACCCTGCT 120

Qy    211 TCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCT 270
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCT 180
```

Db   271 TCCTGACTGGCATAGTCGGCAATGGATTGGTGATCCTCGGTCATGGGTTACCAGAAGAAGC 330
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181 TCCTGACTGGCATAGTCGGCAATGGATTGGTGATCCTCGGTCATGGGTTACCAGAAGAAGC 240

Qy   331 TAAGGAGCCATGACGGACGGACAAGTACCGGCTGCACCTGTCAGTCAGTGGCTGACCTCCTTGTCA 390
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TAAGGAGCCATGACGGATGACGGACAAGTACCGGCTGCACCTGTCAGTCAGTGGCTGACCTCCTTGTCA 300

Qy   391 TCACACTCCCCTTCTGGGCAGTTGATGCCATGGCCATGGCTGGTACTTTGGGAAATTTTGT 450
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TCACACTCCCCTTCTGGGCAGTTGATGCCATGGCCATGGCTGGTACTTTGGGAAATTTTGT 360

Qy   451 GTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCT 510
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCT 420

Qy   511 TCATCAGCCTGACCGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGA 570
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCATCAGCCTGACCGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGA 480

Qy   571 AACTGCTGGCTGAAAAGGCAGTCTATGTGGGGGTCTGGATCCCCAGCCCTCCTCCTGACTA 630
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AACTGCTGGCTGAAAAGGCAGTCTATGTGGGGGTCTGGATCCCCAGCCCTCCTCCTGACTA 540

Qy   631 TACCTGACTTCATCTTTGCCGACGTCAGCCAGGGACATCAGTCAGTGGGGATGACAGGT 690
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TACCTGACTTCATCTTTGCCGACGTCAGCCAGGGACATCAGTCAGTGGGGATGACAGGT 600

Qy   691 ACATCTGTGACCGGCCTTTACCCCGATAGCCTGTGATGGTGGTGTTTCAATTCCAGCATA 750
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ACATCTGTGACCGGCCTTTACCCCGATAGCCTGTGATGGTGGTGTTTCAATTCCAGCATA 660

Qy   751 TAATGTGGGTCTCATCCTGCCGGCATCGTCATCCTCTCCGTGTTACTGCATCATCATCT 810
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TAATGTGGGTCTCATCCTGCCGGCATCGTCATCCTCTCCGTGTTACTGCATCATCATCT 720

Qy   811 CTAAGTCGTCACTCCAAGGCCACCAGAGCGCAAGGCCCTCAAGACGACCAGTCATCC 870
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCAAGTCGTCACTCCAAGGCCACCAGAGCGCAAGGCCCTCAAGACGACCAGTCATCC 780

Qy   871 TCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATATGTGGGGATCAGCATCGACTCCT 930
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATATGTGGGGATCAGCATCGACTCCT 840

Qy   931 TCATCTTTTGGGAGTCATCAAGCAAGGAATGTGACTTCGAGAGCATTGTGCACAAGTGGA 990
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TCATCTTTTGGGAGTCATCAAGCAAGGAATGTGACTTCGAGAGCATTGTGCACAAGTGGA 900

Qy   991 TCTCCATCACAGAGGCCCTTCTTCCTCCACTGTTGCCTGCCTGGAACCCCATCCTGATGCCT 1050
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTCCATCACAGAGGCCCTTCTTCCTCCACTGTTGCCTGCCTGGAACCCCATCCTGATGCCT 960

Qy 1051 TCCTCGGGGCCAAGTCAAAAGCTCTGCCCAGCATGCAGCATCCAACTCAACTCCATGAGCAGAGGCT 1110
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TCCTCGGGGCCAAGTCAAAAGCTCTGCCCAGCATGCAGCATCCAACTCAACTCCATGAGCAGAGGCT 1020

Qy 1111 CCAGGCTCAAGATCCTTCAAAAGGAAAAGCGGGGGTGGACACTCTTCCGTCTTCCACGGAGT 1170
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CCAGGCTCAAGATCCTTCAAAAGGAAAAGCGGGGGTGGACACTCTTCCGTCTTCCACGGAGT 1080

Qy 1171 CAGAATCCTCCAGTTTCACTCCAGTAACCCTTATGCAAGACTTATATACACTT 1230
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 CAGAATCCTCCAGTTTCACTCCAGTAACCCTTATGCAAGACTTATATACACTT 1140

Qy 1231 TATATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAAGACTGACCAGT 1290
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TATATATGATAAAGAACTTTTTTATGTTACTCATTTTCCAGATATAAGACTGACCAGT 1200

Qy 1291 CTTGTACAGTTTTTTTTTTT 1313
         ||||||||||||||||||||
Db 1201 CTTGTACAGTTTTTTTTTTT 1223

RESULT 9
AC122097
LOCUS   AC122097   190149 bp   DNA   linear   HTG 15-NOV-2002
DEFINITION Rattus norvegicus clone CH230-108A24, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC122097
VERSION   AC122097.4 GI:25008263
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 190149)
AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi.V., Aoyagi.A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos.I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila.M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Derano,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,A., Durdin,B., Eaves,K.,
            Egan,A., Escotto.M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,M., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,S., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kova,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill.T.Z., Meenen,E.,
            Milosavljevic.A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarmpunasgoon,A., Pai,S., Parks,K.,
            Pasternak,S., Paul.H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,M., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,A., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson.I., Sitter,C.D., Smajs,D.,
            Sneed,J., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Vaias,A., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams.G., Willson,R., Wilczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,M., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhou,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock.G. and Gibbs,R.A.
TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 190149)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (22-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 190149)
AUTHORS   Rat Genome Sequencing Consortium.

TITLE   Direct Submission
JOURNAL   Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 15, 2002 this sequence version replaced gi:23269619.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            ------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ------- Project Information
            Center project name: GXCR
            Center clone name: CH230-108A24
            ------- Summary Statistics
            Assembly program: Phrap; version 0.990329
            Consensus quality: 179135 bases at least Q40
            Consensus quality: 181484 bases at least Q30
            Consensus quality: 183011 bases at least Q20
            Estimated insert size: 187231; sum-of-contigs estimation
            Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            *   consists of 4 contigs. The true order of the pieces
            *   is not known and their order in this sequence record is
            *   arbitrary. Gaps between the contigs are represented as
            *   runs of N, but the exact sizes of the gaps are unknown.
            *   This record will be updated with the finished sequence
            *   as soon as it is available and the accession number will
            *   be preserved.
            *   1   40067: contig of 40067 bp in length
            *   40068   40167: gap of unknown length
            *   40168 114660: contig of 74493 bp in length
            *   114661 114760: gap of unknown length
            *   114761 188186: contig of 73426 bp in length
            *   188187 188286: gap of unknown length
            *   188287 190149: contig of 1863 bp in length.
FEATURES   Location/Qualifiers
     source   1..190149
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-108A24"
ORIGIN

Query Match   62.3%;   Score 1169.4; DB 2;   Length 190149;
Best Local Similarity 84.1%;   Pred. No. 2.9e-232;
Matches 1465; Conservative 0; Mismatches 156; Indels 122; Gaps 8;

Qy   119 CATGGAACCGATCAGTCAGTGTGAGTATATACACTTCGATAATACTACTCCGAGAAGTGGGGTC 178
           ||||||| |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   20640 CCTGTCCCTGTTTCTTTTGCAGATATACACTTCCGATAATACTACTCCGAGAAGTGGGGTC 20699

Qy   179 TGGAGCTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTCAATAG 238
           ||||||||||||||||||||||||||||||| |||||||||||||| ||||||| ||
Db   20700 TGGAGCTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGAAAACTTCAACAG 20759

Qy   239 GATCTTCCTGCCCACCATCTACTTCATCATCTTCATCATCTTCTTGACTGGCAATGGATT 298
           ||||||||||||||||||||||||||||||| ||||||| |||||||||||||||||||
Db   20760 GATCTTCCTGCCCACCATCCATCATTTATCATCTTCTTGACTGGCATAGTGGCAATGGGTT 20819

```
Qy    299 GGTGATCCTGGTCATGGGTTACCAGAAGAGCTAAGGAGCATGACGGACAAGTACCGGCT 358
          ||| ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  20820 GGTAATCCTGGTCATGGGTTACCAGAAGAGCTGAGGAGCATGACGGACAAGTACCGGCT 20879

Qy    359 GCACCTGTCAGTGGCTGACCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGC 418
          |||||||| ||||||||||||||||||||||||||||| ||||||||||||||||||
Db  20880 GCACCTGTCCGTGGCTGACCTCCTCGTTGTCATCACACTCCCCTTCTGGGCAGTGACGC 20939

Qy    419 CATGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTCACACTGT 478
          ||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||| |
Db  20940 CATGGCTGACTGGTACTTTGGGAAATTTTATGTAAGGCTGTCCATATCATCTACACCGT 20999

Qy    479 CAACCTTACAGAGACGGTTCTCATCCTGGCCTTCATCAGCCTGGACCTGTACCTCGCCAT 538
          ||| ||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  21000 CAACCTTACAGCAGTGTTCTCATCCTGCCCTTCATCAGCCTGGACCTGCACCTTGCCAT 21059

Qy    539 TGTCCAGGCCACCACAGTCAAAGGCTGCGGCTGAAAAGGCAGTCTATGT 598
          ||||||||||||||| |||||||||||||||||||||||||||||||||
Db  21060 TGTCCAGGCCACCACAGTCAAAGGCCGCGGCTGAAAAGGCCGTCTATGT 21119

Qy    599 GGGCGCTCTGGATCCCAGCGCCTCCCTGACTATACCTTGCATCTTGCCCACGTCAG 658
          ||| ||||||||||||||||| |||||||||||| |||||||||||||||||||||
Db  21120 GGGTGTCTGGATCCCCAGCGCCTCCCTGACTATCCTCGCCGCTAGTCAG 21179

Qy    659 CCAGGGGACATCAGTCAGGGGGATGACAGGTACATTGTGACCGCCTTTACCCCGATAG 718
          |||||||||                     ||||||||||||||||| ||||||||||
Db  21180 CCAGGGGAC---------GGCAGTACATCTGTGACCGCCTTTACCCCGACAG 21224

Qy    719 CCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCATCCTGCCGGGAT 778
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  21225 CCTGTGGATGGTGCTGTTCCAGCACACATCGGTGGGTCATCTCTGCCGGGCAT 21284

Qy    779 CGTCATCCTCTCCGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCACCA 838
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  21285 CGTCATCCTCTGTCCGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCACCA 21344

Qy    839 GAAGCGCAAGGCCCTCAAGACGACCAGTCATCATCCTAGCTTTCTTTGCCTGCTGGCT 898
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  21345 GAAGCGCAAGGCCCTCAAGACTACCGGTCATCCTTATCCTGGCTTTCTTTGCCTGCTGGCT 21404

Qy    899 GCCATATTATGTGGGGGATCAGCATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGG 958
          || |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21405 ACCGTATTACGTGGGGGATCAGCATCAGCATCGATTCCTTCATCCTTTGGAGTCATCAAGCAAGG 21464

Qy    959 ATGTGACTTCGAGAGCATTGTGACAAGTGGATCTCCATCACAGAGAGCCCTCGCCTTCTT 1018
          ||| ||||||| ||||||||||||||||||||||||||||||||| |||||||||||||
Db  21465 ATGTGAGTTCGAGAGCGTCGTGTGCAAGTGGATCTCCATCCCATCACCGGAGGCCCTCGCCTTCTT 21524

Qy   1019 CCACTGTTGCCTGAACCCCATCCTATGCCTTCCTCGGGGCCAAGTCAAAAGCTCGC 1078
          |||||| ||||||||||||||| |||||||||||||||||||||| ||||||||||
Db  21525 CCACTGTTGCCTGAACCCCATCCTCTACGCCTTCCTCGGGGCCCAAATTCAAGAGCTCCGC 21584

Qy   1079 CCAGCATGCACTCAACTCAATCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAA 1138
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  21585 GCAGCATGCACTCAACTCAATCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGGAA 21644

Qy   1139 GCGGGGGTGGACACTCCTTCCGTCCCACCGAGTCAGAATCCTCCAGTTCACTCCAGCTA 1198
          ||||| ||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  21645 ACGGGGGTGGACACTCCTTCCGTCTCCACAGAGTCAGAATCCTCCAGTTTCACTCCAGCTA 21704

Qy   1199 ACCCTTATGCAAAGACTTATATATATATATAT--------GATAAAGAACTT 1249
          ||| ||||| ||||||||||||||||||||||        | |||||||||||
Db  21705 ACACTTACGCAAAGACACATATATATATATATATATATATATGACAAAGAACTTT 21764

Qy   1250 TTTTATGTTACACATTTCCAGATATAGAAGACTGACCAGTCTTGTACAGTTTTTTTTTTT 1309
          ||||||||||||||||| ||||||||||||| ||||||||||||||||||||||||||||
Db  21765 TTTTATGTTACACATTTCCAGATATAAAAGAACTGACCAGTTTGTTCAGTTTTTTTTTTTAT 21824

Qy   1310 TTTTAATTGACTGTTGGGAGTTTATGTTCCTCT------AGTTTTTGTGAGGTTTGAC 1362
          |||| ||||||||||||||||||||||||||||      ||||||||||||||||||||
Db  21825 TGACTGTTGGGAGTTTATGTTTTCTCTAGTTTTTTGTGTGTGTGTGTGTGAGGTTTGAC 21884
```

```
Qy   1363 TTAATTATATAAATATTGTTTTT--------GTTGTTGTTCATGTGAATGAGCGTC 1411
          |||||||||||||||||||||||        ||||| || || ||||||||||||||
Db  21885 TTAATTATATAAATATTGGTTTTTTCTGTTGTTGTTGTCTGTTCACGTGAATGAGTGTC 21944

Qy   1412 TAGGCAGGACCTGTGGCCAAGTCTTAGTAGCTGTTTATCTGTGTGTAGAGGACTGTAGAAC 1471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21945 TAGGCAGGACCTGTGGCCAAGTCTTAGTAGCTGTTTATCTGTGTGTAGAGGACTGTAGAGC 22004

Qy   1472 TGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGTCTAGCCGTGA 1531
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  22005 TGTAGAGGAAGAAGAACTGAACGCTCCAGAATGTGTGGTAAATCGGATAAAGTCTAGAGTGA 22064

Qy   1532 TCCTCAGCTGTCTGCATAATCTCTTCATTCCGAGAGCACCCCACCCCCACCCACC 1591
          ||||||||||| |||||||||||||||||                       |||
Db  22065 CCCTCAGCTGTCGCGCTCCATAATCTCTT--------------------------CGCCCCG 22098

Qy   1592 CCCACCCCATCTTAATTGTTGGTTATGCTGTGTGATGGTTTGTTTGGTTTTTTTTTTTG 1651
          |||||||| |||||||| ||||||||||||||||||||                      
Db  22099 CCCACCCCGTTCTGAATTGTTTGGTTGTGCGTGTG-------------------------- 22135

Qy   1652 TTGTTGTTGTTGTTGTTTTTTTCTGTAAAGATGGCACTTAAAAACCAAAGGCTGAAATGG 1711
                                  |||||||||||||||||||||||||||||||||||
Db  22136 ------------------------GTTTCCTCGTAGAAGATGGCACTTAAAACCAAAGGCTGAAATGG 22179

Qy   1712 TGGTAGAAATGCTGGGGGTTTTTGTTTTGTTTTTCTTTCAAGAGTAGATTG 1771
          |||||||||||||                       ||||||||||||||||||
Db  22180 T-ATAGAAATGCTGG-------------------GTTTTTTCTTTCCAGTTTTCAAGGGTGGATTG 22224

Qy   1772 ACTTCAGTCCCTACAAATGTACAGTCTTGATTACATTGTAATAAAGTCAATGATAAA 1831
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  22225 ATTTCAGTACCTACAAATGTACAGTCTTGTATTATGTGTGTTAATAAAAGTCAATGTAAA 22284

Qy   1832 CTT 1834
          |||
Db  22285 CTT 22287
```

```
RESULT 10
MMLCR13
LOCUS       MMLCR13                  1422 bp    DNA     linear   ROD 31-OCT-1996
DEFINITION  Mus musculus lcr-1 gene, exon 2.
ACCESSION   Z80113
VERSION     Z80113.1  GI:1655638
KEYWORDS    CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1422)
  AUTHORS   Schubel,A., Burgstahler,R. and Lipp,M.
  TITLE     The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
            High expression in thymus and lymphoid tissues
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1422)
  AUTHORS   Lipp,M.
  TITLE     Direct Submision
  JOURNAL   Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
            Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
            D-13122, GERMANY
FEATURES             Location/Qualifiers
     source          1..1422
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV"
                     /db_xref="taxon:10090"
                     /clone_lib="lambda fix (TM) II"
                     /dev_stage="adult"
     intron          <1..259
                     /number=1
     gene            260..1422
                     /gene="lcr-1"
     exon            260..>1422
                     /gene="lcr-1"
```

```
Qy  1081 AGCATGCACTCAACTCCATGAGCAGAGAGGCTCCAGCTCCAGATCCTTCCAAAGGAAAGC 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  1200 AGCATGCACTCAACTCCATGAGCAGAGAGGCTCCAGCTCCAGATCCTTTCCAAAGGAAAGC 1259

Qy  1141 GGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAAC 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260 GGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAAC 1319

Qy  1201 CCTTATGCAAAGACTTATATATATATATATATATATAAGAACTTTTTATGTTAC 1260
         |||||||||||||||||||||||||||||||||       |||||||||||||
Db  1320 CCTTATGCAAAGACTTATATATATATATATATATATTGATAAAGACTTTTTATGTTAC 1379

Qy  1261 ACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT 1303
         |||||||||||||||||||||||||||||||||||||||||||
Db  1380 ACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT 1422
```

RESULT 11
MMLCR12
LOCUS   MMLCR12   1180 bp   mRNA   linear   ROD 18-NOV-1997
DEFINITION Mus musculus lcr-1 gene.
ACCESSION   Z80112
VERSION   Z80112.1 GI:2632100
KEYWORDS   CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1180)
AUTHORS   Schubel,A., Burgstahler,R. and Lipp,M.
TITLE   The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
            High expression in thymus and lymphoid tissues
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1180)
AUTHORS   Lipp,M.
TITLE   Direct Submission
JOURNAL   Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
            Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
            D-13122, GERMANY
COMMENT   On Nov 21, 1997 this sequence version replaced gi:1542888.
FEATURES   Location/Qualifiers
     source   1..1180
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129SV"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
                     /dev_stage="adult"
     gene   1..1180
                     /gene="lcr-1"
     CDS   1..1074
                     /gene="lcr-1"
                     /function="receptor for chemokine SDF-1"
                     /standard_name="CXC chemokine receptor 4"
                     /codon_start=1
                     /product="CXCR-4"
                     /protein_id="CAB02202.1"
                     /db_xref="GI:1542889"
                     /db_xref="GOA:P70658"
                     /db_xref="SWISS-PROT:P70658"
                     /translation="MEPISIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTI
                     YFIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVIL.PFWAVDAMADW
                     YFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGV
                     WIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLMVVFQFQHIMVGLILPGI
                     VILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISDSFILLGVIK
                     QGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKIL
                     SKSKRGGHSSVSTESESSSFHSS"
     exon   <1..15
                     /gene="lcr-1"
                     /number=1
     exon   16..>1180
                     /gene="lcr-1"
```

```
                     /product="CXCR-4"
                     /standard_name="CXC chemokine receptor 4"
                     /number=2
                     /function="receptor for chemokine SDF-1"

ORIGIN

Query Match      61.4%;   Score 1152.6;  DB 10;  Length 1422;
Best Local Similarity  98.4%;  Pred. No. 4.9e-229;
Matches 1164;  Conservative  0;  Mismatches  19;  Indels  0;  Gaps  0;

Qy   121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG 180
         ||||||||      ||||||||||||||||||||||||||||||||||||||||||||||
Db   240 TGTCCCGTTTCCTTTGCAGATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG 299

Qy   181 GAGACTATGACTCCAACAAGGAACCCGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   300 GAGACTATGACTCCAACAAGGAACCCGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 359

Qy   241 TCTTCCTGCCCACCATCACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 300
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   360 TCTTCCTGCCCACCATCATCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 419

Qy   301 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGACGCATGACGGACAAGTACCGGCTGC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   420 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGACGCATGACGGACAAGTACCGGCTGC 479

Qy   361 ACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCA 420
         |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   480 ACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCA 539

Qy   421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   540 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCACACTGTCA 599

Qy   481 ACCTCACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTG 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600 ACCTCACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTG 659

Qy   541 TCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660 TCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 719

Qy   601 GCGTCTTGATCCCAGCCCTCCTGCTCACTATACCTGACTTCATCTTTGCCGACGTCAGCC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720 GCGTCTTGATCCCAGCCCTCCTGCTCACTATACCTGACTTCATCTTTGCCGACGTCAGCC 779

Qy   661 AGGGGACATCAGTCAGGGATGACAGCAGGTACATCTGTGACCGCCTTACCCGATAGCC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780 AGGGGACATCAGTCAGGGATGACAGCAGGTACATCTGTGACCGCCTTTACCCGATAGCC 839

Qy   721 TGTGGATGGTGTGTTTCAATTCCAGCATATAAATGGGGTCTCATCCTGCCGGGATCG 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   840 TGTGGATGGTGTGTTTCAATTCCAGCATATAAATGGGGTCTCATCCTGCCGGGCATCG 899

Qy   781 TCATCCTCTCTGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCACCAGA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 TCATCCTCTCTGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCACCAGA 959

Qy   841 AGCGCAAGGCCCTCAAGACAGCAGTGGATCCTCATCACAGAGGCCCTGCCTGCTGC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 AGCGCAAGGCCCTCAAGACGACAGTGGATCCTCATCACAGAGGCCCTGCCTGCTGC 1019

Qy   901 CATATATGTGGGATCCAGCATCAGACTCCTTTGGGAGTCATCAAGCAAGGAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 CATATTATGTGGGATCAGCATCGACTCCTTTGGGAGTCATCAAGCAAGGAT 1079

Qy   961 GTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1080 GTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCC 1139

Qy  1021 ACTGTTGCCTGAACCCCATCCTTATGCCTTCCTCGGGGCCCAAGTTCAAAAGCTCGCCC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1140 ACTGTTGCCTGAACCCCATCCTTATGCCTTCCTCGGGGCCCAAGTTCAAAAGCTCGCCC 1199
```

```
ORIGIN          /number=2

Query Match         60.9%;   Score 1143.2;  DB 10;  Length 1180;
Best Local Similarity   99.1%;   Pred. No. 4.3e-227;
Matches 1160;   Mismatches  8;   Indels  2;   Gaps  1;

Qy    136  TGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGAGACTCCA  195
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11  TCAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGAGACTCCA  70

Qy    196  ACAAGGAACCCTGCTTCCGGGATGAAAAGTCCATTTCAATAGGATCTTCCTGCCCACCA  255
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71  ACAAGGAACCCTGCTTCCGGGATGAAAAGTCCATTTCAATAGGATCTTCCTGCCCACCA  130

Qy    256  TCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGG  315
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131  TCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGG  190

Qy    316  GTTACCAGAGAAGAGAAGCTAAGGACATGAGCGGCAGCAAGTACCGGCTGCACCTGTCAGTGGCGTG  375
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191  GTTACCAGAGAAGAGAAGCTAAGGACATGAGCGGCAGCAAGTACCGGCTGCACCTGTCAGTGGCGTG  250

Qy    376  ACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGGTGGTTACT  435
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251  ACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGGTGGTTACT  310

Qy    436  TTGGGAAATTTTGTGTGTAAGGCTGTCCAATATCATCTCAACCTGTCAACCCTACAGCAGCG  495
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311  TTGGGAAATTTTGTGTGTAAGGCTGTCCAATATCATCTCAACCTGTCAACCCTACAGCAGCG  370

Qy    496  TTCTCATCCTGCCTTCATCAGCCTTGACACCGGTACCTCGGCCATTGTCCACGCCACCAACA  555
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371  TTCTCATCCTGGCCTTCATCAGCCTTGACACCGGTACCTCGGCCATTGTCCACGCCACCAACA  430

Qy    556  GTCAAAGGCCAAGGAACGAAACTGCTACGCAGTCATATGTGGGCGTCTGAATCCCAG  615
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431  GTCAAAGGCCAAGGAACTGCTGGCTGGGCTGAAAAGGCAGTCATGTGGGCGTCTGGATCCCAG  490

Qy    616  CCCTCCTCCTGACTATACCCGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTC  675
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491  CCCTCCTCCTGACTATACCCGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTC  550

Qy    676  AGGGGGATGACAGGTACAGGTACATCTGTGACATCTGGTCTCATGTGGGGCATCCTCCTGTGT  735
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551  AGGGGGATGACAGGTACAGGTACATCTGTGACATCTGGTCTCATGTGGGGCATCCTCCTGTGT  610

Qy    736  TTCAATTCCAGCATATAATGGTGGTCTCACACTTCATCTCCTGCCCGGCATCGTCATCCTCTCCTGTT  795
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    611  TTCAATTCCAGCATATAATGTGGGTCTCACACTTCATCTCCTGCCCGGCATCGTCATCCTCTCCTGTT  670

Qy    796  ACTGCATCATCATCCTCTAAGCTGTCTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCA  855
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    671  ACTGCATCATCATCCTCTAAGCTGTCTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCA  730

Qy    856  AGAGCAACAGCAGTCATCCTCATCCTCTAGCTTCTTGCCTGGCTGCTGCCATATTATGTGGGGA  915
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    731  AGACGACAGCAGTCATCCTCATCCTCTAGCTTTCTTGCCTGGCTGCTGCCATATTATGTGGGGA  790

Qy    916  TCAGCATCGACTCCTTTTGGGAGTCATCAAGGATCATCAAGCAAGGATGTGACTTCGAGAGCA  975
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    791  TCAGCATCGACTCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCA  850

Qy    976  TTGTGCACAAGTGGATCTCCATCACAGAGGGCCCTCGGCCTTCTTGTTGCCTCGAACC  1035
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    851  TTGTGCACAAGTGGATCTCCATCACAGAGGGCCCTCGGCCTTCTTGTTGCCTCTGAACC  910

Qy   1036  CCATCCCTATGCCTTCCTCGGGGCCAAGTCAAAAGTCTTGCCCAGCATGCACTCAACT  1095
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    911  CCATCCCTATGCCTTCCTCGGGGCCCAAGTCAAAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACT  970

Qy   1096  CCATGAGCAGAGGCTCCAGCCTCCAAGATCCTTCCAAAGAAAGCGGGGTGACACTCTT  1155
                |||||||||
Db    971  CCGTTCCACGGAGTCAGAATCCTCCAGTTTCACTCAGTTTCACTCCAGCTCACCCTTATGTTCACTCCTTATGCAAAGACT  1030

Qy   1156  CCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCAGTTTCACTCCAGCTCACCCTTATGTTCAC--ATTTTCCAGAT  1215
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1031  CCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCAGTTTCACTCCAGCTCACCCTTATGTCAAAGACT  1090

Qy   1216  TATATAATATATATAATATATATGATAAAGAACTTTTTTATGTTACAC--ATTTTCCAGAT  1273
                ||||||||||||||||||||||||||||||||||||||||||||||
Db   1091  TATATAATATATATATAATATATTGATAAAGACTTTTTTATGTTACCATTTTTCCAGAT  1150

Qy   1274  ATAAGAGACTGACCAGTCTTGTACAGTTTT  1303
                ||||||||||||||||||||||||||||
Db   1151  ATAAGAGACTGACCAGTCTTGTACAGTTTT  1180
```

```
RESULT 12
AC107097/c
LOCUS        AC107097           240931 bp    DNA       linear   HTG 13-MAY-2003
DEFINITION   Rattus norvegicus clone CH230-175O10, *** SEQUENCING IN PROGRESS
             ***.
ACCESSION    AC107097
VERSION      AC107097.5  GI:30580624
KEYWORDS     HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE       Rattus norvegicus (Norway rat)
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE    1  (bases 1 to 240931)
  AUTHORS    Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
             Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
             Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
             Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
             Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
             Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
             Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
             Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
             Cleveland,C., Cockrell,R., Cox.C., Coyle,M., Cree,A., D'Souza,L.,
             Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
             Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
             Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
             Egan,A., Escotto,M., Eugene,C., Evans,C., Falls,T., Fan,G.,
             Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
             Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
             Gebregeorgis,E., Geer,K., Gill,R., Grady,W., Guevara,W., Guevara,W.,
             Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
             Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
             Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
             Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
             Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,M., Jolivet,A.,
             Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
             Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
             Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
             Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
             Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
             Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
             Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
             Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
             Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
             Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
             Nwaokelemeh,O., Okwuonu,G., Olarmpunsagoon,A., Pal,S., Parks,K.,
             Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
             Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
             Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
             Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
             Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.,
             Sanders,W., Savery,G., Scherer,S., Scott,S., Shatsman,S., Shen,H.,
             Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
             Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,D.,
             Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
             Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,S., Usmani,K.,
             Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
             Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
```

```
                Williams,G., Willson,R., Wlecyzk,R., Wooden,H., Worley,K.,
                Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
                Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
                Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
                Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 240931)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-2002) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 240931)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 13, 2003 this sequence version replaced gi:23115278.
                The sequence in this assembly is a combination of BAC based reads
                and whole genome shotgun sequencing reads assembled using Atlas
                (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
                in the feature table below represents a scaffold in the Atlas
                assembly (a 'contig-scaffold'). Within each contig-scaffold,
                individual sequence contigs are ordered and oriented, and separated
                by sized gaps filled with Ns to the estimated size. The sequence
                may extend beyond the ends of the clone and there may be sequence
                contigs within a contig-scaffold that consist entirely of whole
                genome shotgun sequence reads. Both end sequences and whole genome
                shotgun sequence only contigs will be indicated in the feature
                table.
                ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
                ---------- Project Information
                Center project name: GKBK
                Center clone name: CH230-175O10
                ---------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 216793 bases at least Q40
                Consensus quality: 219254 bases at least Q30
                Consensus quality: 221224 bases at least Q20
                Estimated insert size: 228277; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
                ----------
                * NOTE: Estimated insert size may differ from sequence length
                * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
                * NOTE: This is a 'working draft' sequence. It currently
                * consists of 1 contigs. Gaps between the contigs
                * are represented as runs of N. The order of the pieces
                * is believed to be correct as given, however the sizes
                * of the gaps between them are based on estimates that have
                * provided by the submittor.
                * This sequence will be replaced
                * by the finished sequence as soon as it is available and
                * the accession number will be preserved.
FEATURES            Location/Qualifiers
     source         1..240931
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-175O10"
     misc_feature   1..1249
                    /note="wgs end_extension
                    clone_end:Sp6"
     misc_feature   complement(5161..5587)
                    /note="clone_boundary
                    clone_end:Sp6
                    site:EcoRI
                    end_sequence:BH357332"
     misc_feature   complement(232579..233220)
                    /note="clone_boundary
                    clone_end:T7"
                    site:EcoRI
                    end_sequence:BH357330"
     misc_feature   235468..237558
                    /note="wgs end_extension
                    clone_end:T7"
     misc_feature   239450..240931
                    /note="wgs end_extension
                    clone_end:T7"
ORIGIN

Query Match                59.4%;   Score 1114;  DB 2;  Length 240931;
Best Local Similarity      84.2%;   Pred. No. 9.8e-221;
Matches 1400;  Conservative   0;  Mismatches 140;  Indels 122;  Gaps  8;

Qy     200 GGAACCCGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTA 259
           |||||||||||||||||||||||   ||||||||| ||||||||||||||||||||||||
Db  237558 GGAACCCGCTTCCGGGATGAAAACTTCAACAGGATCTTCCTGCCCACCATCTA 237499

Qy     260 CTTCATCATCTTCTTGACTGGCATAGTCGGCATAGTCGGCAATGGATCGGTGATCGTGGTCATGGGTTA 319
           |||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||||||||
Db  237498 TTTTATCATCTTCTTGACTGGCATAGTCGGCATAGTCGGCAATGGGTTGGTAATCCTGGTCATGGGTTA 237439

Qy     320 CCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCGTGCACCTGTCAGTGGCTGACCT 379
           |||||||||||| |||||||||| |||| |||||||||||| |||||||||| ||||||||
Db  237438 CCAGAAGAAGCTGAGGAGCATGACAGACAGACAAGTACCGGCTGCACCTGTCCGTGGCTGACCT 237379

Qy     380 CCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCCATGGCTGACTGGTACTTTGG 439
           ||||| ||||||||||||||||||||||||||||| | |||||||||||||||||||||||||||
Db  237378 CCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGGACCCATGGCCATGGCTGACTGGTACTTTGG 237319

Qy     440 GAAATTTTGTGTAAGGCTGTCCATATCATCATCATCTACACTGTCAACCTCTCACAGCAGCGTTCT 499
           |||||||| |||||||||||| |||||||||||||| ||||| ||||||||  |||||||||||| 
Db  237318 GAAATTTTATGTAAGGCTGTGCATATCATCATCATCACACGTCAACCTTTACAGCAGTGTTCT 237259

Qy     500 CATCCTGGCCTTCATCAGCTGACACGGTACCTGCCATTGTCACGGCACCAACAGTCA 559
           ||||||||||||||||||||||||||||||| |||||||| ||||||||||||| ||
Db  237258 CATCCTGGCCTTCATCAGCTGACACGGTACCTTGCCATTGTCCACGGCACCAACAGCCA 237199

Qy     560 AAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCT 619
           |||||| ||||| |||||||||||||||||| ||||||||||||| || ||||| || ||
Db  237198 GAGGCCGAGGAAGTGCTGGCTGAAAAGGCCGTCTATGTGGGCGTGTCTCGGATCCCGCCCCT 237139

Qy     620 CCTCCTGACTATACCTGACTTCATCTTTGCCGACATGTCAGCCAGGGGGACATCAGTCAGGG 679
           ||||||||||| ||||||| ||||||| ||||||||||| |||||||||||           
Db  237138 CCTCCTGACTATCCCTGACACATCATCTTCGCCGACATGTCGCCAGGGGGAC--------- 237090

Qy     680 GGATGACAAGTACATCTGTGACCGGCCTTACCCTGGTGGATGGTGGTGTTCA 739
               |||| ||||||||||||||||||||||||||||||||| |||||||| ||
Db  237089 ----GGCAGTGACATCTGTGACCGGCCTTACCCTGGTGGATGGTGTGTTCCA 237034

Qy     740 ATTCCAGCATATAATGGTGGTCTCATCCTGCCCGGCATCGTCATCCTCGTCCTGTTACTG 799
            |||||||| || ||||||| |||||||||||||||||||||||||||||||||||||||
Db  237033 GTTCCAGCACATCATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCGTCCTGTTACTG 236974

Qy     800 CATCATCATCTCAAGCTGTCCACACTCCAAGGGCCACCAGAGAGGCAAGGCCCTCAAGAC 859
           |||||||||||| ||||||||||||||||||||||||||||| || |||||||||||||
Db  236973 CATCATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAGCGCAAGAGCCCTCAAGAC 236914

Qy     860 GACAGTGCATCCTCATCCTCTAGGTTTCTTTGCCTGCTGCGGCTGCCCATATTATGTGGGGATCAG 919
              |||| |||||| ||||| | |||||||||||||||||   ||| ||| || ||||||||||||||
Db  236913 TACGGTCATCCTTATCCTTGGCTTTCTTTGCCTGCTGGCTACCGGTATTACGTGGGGATCAG 236854

Qy     920 CATCGGACTCCTTCATCCTTTTGGGAGTCATCAGCAAGCAAGGATGGACTTCGAGAGCATTGT 979
           ||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||| | ||
Db  236853 CATCGGATTCCTTCATCCTTTTGGAGGTCATCAGCAAGCAAGGATGGACTTCGAGAGCGTCGT 236794

Qy     980 GCACAAGTCGGATCTCCATCACAGAGGCCCCTGCCTTCTTCCACTGTTGCCCACTGTTGCCCAATCCCCAT 1039
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||| |||||
Db  236793 GCACAAGTCGGATCTCCATCACGGAGGCCCCTGCCTTCTTCCACTGTTGCCCACTGTTGCCCAACCCCCAT 236734
```

```
JOURNAL     J. Immunol. 157 (12), 5455-5460 (1996)
MEDLINE     97113334
PUBMED      8955194
REFERENCE   2  (bases 1 to 3366)
AUTHORS     Heesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
TITLE       Direct Submission
JOURNAL     Submitted (30-JUL-1996) Pathology, Harvard Medical School, 200
            Longwood Avenue, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..3366
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
     gene            1..3366
                     /gene="CXCR-4"
     CDS             join(1..21,2308..3366)
                     /gene="CXCR-4"
                     /note="SDF-1 chemokine receptor"
                     /codon_start=1
                     /product="fusin"
                     /protein_id="AAC52953.1"
                     /db_xref="GI:3645915"
                     /translation="MEPISVSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDRYICDRLYPDSLWMVFQFQHIMVGLILP
                     GIVILSCYCIIISKSHSKGHQKRALKTVILLAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGRGGHSSVSTSESSSFHSS"
     exon            1..21
                     /gene="CXCR-4"
                     /number=1
     exon            1..15
                     /gene="CXCR-4"
                     /note="alternate exon 1"
     intron          16..>2307
                     /gene="CXCR-4"
                     /note="alternate intron 1"
     intron          22..2307
                     /gene="CXCR-4"
                     /number=1
     exon            2308..3366
                     /gene="CXCR-4"
                     /number=2
ORIGIN

Query Match        56.2%;    Score 1055;    DB 10;    Length 3366;
Best Local Similarity 98.6%;    Pred. No. 1.1e-208;
Matches 1064;    Conservative 0;    Mismatches 15;    Indels 0;    Gaps 0;
```

```
Qy     1040  CCTCTATGCTTCCTTCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCAT  1099
             |||||  || |||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db   236733  CCTCTACGCCTTCCTTCTCGGGGCCCAAATTCAAGAGCTCCGCGCAGCATGCACTCAACTCCAT  236674

Qy     1100  GAGCAGAGAGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGT  1159
             |||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db   236673  GAGCAGAGAGCTCCAGCCTCAAGATCCTTCCAAAGGGAAACGGGGTGGACACTCTTCCGT  236614

Qy     1160  CTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTATGCAAAGACTTATA  1219
             |||| ||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   236613  CTCCACAGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACACTTACGCAAAGACATATA  236554

Qy     1220  TAATATATATATATAT-------GATAAAGAACTTTTTATGTTACACATTTTCCA  1270
             |||||||||||||||||       |||||||||||||||||||||||||||  |||
Db   236553  TAATATATATATATATATATATATGACAAAGAACTTTTTTTTTTATGTTACACATTTCCCA  236494

Qy     1271  GATATAAGAGACTGACCCAGTCTCGTACACAGTTTTTTTTTTTAATTGACTGTTGGGAG  1330
             |||||||||  ||||||||||||||  |||||||||||||||||||||||||||||||||
Db   236493  GATATAAAAGACTGACCCAATCTCGTACACAGTTTTTTTTTTTATTGACTGTTGGAGTTTATGTTT  236434

Qy     1331  TTTATGTTCCTCT-------AGTTTTGTGAGGTTGACTAATTTATATATAATAT---  1379
             |||||||||||||       ||||||||||||||||||||||||||||||||||||||
Db   236433  CTCTAGTTTTTTGTGTGTGTGTGTGTGTGGTTTGACTTGACTTAATTTATATAAATATTGGT  236374

Qy     1380  -------TGTTTTTGTTGTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGGCCAAG  1432
                    ||||||||||||||||||||||||||||  ||||||  ||||||||||||||
Db   236373  TTTTTCTGTTGTTGTCTTTCACGTGAATGAGTGTCTAGGCAGGACCTGTGGCCAAG  236314

Qy     1433  TTCTTAGTAGCTGTTTATCTGTGTGTGTAGGACTGTAGAACTGTAGAGGAAAACTGAACA  1492
             |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   236313  TTCTTAGTCGCTGTCGTGTCGTGTGTGTAGGACTGTAGAGCTGTAGAGGAAGGAACTGAACG  236254

Qy     1493  TTCCAGAATGTGTGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCATAA  1552
             ||||||||||||||||  |||||||||||||||||  |||||||||||||||||||||||
Db   236253  CTCCAGAAGTGTGTGTAAATCGGATAAAGCTAGAAGTGACCCCTCAGCTGTCGCGCATAA  236194

Qy     1553  TCTCCTCCATTCCGAGGAGCACCCCACCCCCCATTCTTAAAATGT  1612
             |||||||||||||||||||||||||  ||  |||||||||||||||
Db   236193  TCTCTT----------------------CGCCCCGCCACCCCGTTTCCTGAATTGT  236160

Qy     1613  TTGGTTATGCTGTGATGGTTTGGTTTGTTGTGTTTTTTTTTT  1672
             |||||||||||||||  ||||||||||||||||||
Db   236159  TTGGTTGTGCTGTGTG----------------------GTTTC  236139

Qy     1673  TCTGTAAAAGATGGGACTTAAAAACCAAAGCCTGAAATGTGGTGTAGAAATGCTGGGGGTTTT  1732
             ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   236138  TCTGTAGAAGATGGGACTTAAAAACCAAAGCCTGAAATGT-ATAGAAATGCTGG-------  236086

Qy     1733  TTTGTTGTTTGTTTTTCAGTTTTCAAGGATGATTGACTTCCTACACAAATGTA  1792
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   236085  -------GTTTTTTCTTTCAGTTTTCAAGGGTGGATTGATTTCAGTACCTACAAAATGTA  236034

Qy     1793  CAGTCCTTGATTACATTGTTAATAAAAGTCAATGAAATAAACTT  1834
             ||||||||||||||||||||||||||||||||||||||||||||
Db   236033  CAGTCCTTGATTACATTGTTGTTAATAAAAGTCAATGAAATAAACTT  235992
```

```
RESULT 13
LOCUS       MMU65580
MMU65580    Mus musculus fusin (CXCR-4) gene, complete cds.
DEFINITION  Mus musculus fusin (CXCR-4) gene, complete cds.
ACCESSION   U65580
VERSION     U65580.1    GI:1731650
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE   1  (bases 1 to 3366)
  AUTHORS   Heesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
  TITLE     Cloning of the mouse fusin gene, homologue to a human HIV-1
            co-factor
```

```
                    3366 bp    DNA    linear    ROD 13-DEC-1996
```

Query Match
Best Local Similarity
Matches 1064;

```
Qy     121  TGGAACCCATCAGTGCGTGAGTAGTATACACTTCGATAACTACACTACTGAAGAAGTGGGGGTCTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2288  TGTCCCGTGTTTCCTTTGCAGATAGACACTTCTGATAACTACACTACTCGATAACTACACTCTGAAGAAGTGGGGTCTG  2347

Qy     181  GAGACTATGAGCTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGA  240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2348  GAGACTATGAGCTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGA  2407

Qy     241  TCTTCCTGCCCACCATCACTTACTTCATATCTTCTTGACTGAGCATAGTCGGCAATGGATTGG  300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2408  TCTTCCTGCCCACCATCACTTACTTCATATCTTCTTGACTGAGCATAGTCGGCAATGGATTGG  2467

Qy     301  TGATCCTGGTCATGGTTTACCAGAAGAAGCTAAGGAGCGTAAGGAGCAAGTACCGGCTGC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2468  TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCGTAAGGAGCAAGTACCGGCTGC  2527

Qy     361  ACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCTTCTGGCCAGTTGATGCCA  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2528  ACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCTTCTGGCCAGTTGATGCCA  2587

Qy     421  TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAATCACACTGTCA  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Matches 1391;  Conservative 0;  Mismatches 301;  Indels 136;  Gaps 10;

Db   2588 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA  2647

Qy     50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGTGCTGTCCGGTAACCACCGCGTGTAGAGC    109
Db     20 GCGCGGCGCAAAGTGACGCCGAGGCCTGAGTGTGCTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy    110 GAGTGTTGCCATGGACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA   169
Db     80 CAGCGCGGTACCAGGA----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA  134

Qy    170 AGTGGGGTCTGAGACATGACTGCAACAAGGAACCTGCTTCCGGATGAAGAAAGTCCA    229
Db    135 AATGGGCTCAGGGGACTATGACTTTCCGTGAAGAAAATGCTAA                 194

Qy    230 TTTCAATAGGATCTTCCTGCCACCATCACTTCTTCTCTTGACTGGCATAGTCGG      289
Db    195 TTTCAATAAAATCTTCCTGCCCACCATCTCTTCTTAACTGGCATTGTGGG          254

Qy    290 CAATGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAA  349
Db    255 CAATGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGACGGACAA  314

Qy    350 GTACCGGCTGCACCTGCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTGGGC   409
Db    315 GTACAGGCTGCACCTGCAGTGGCCGCACCTCCTCTTGTCATCACGCTTCCCTTCTGGGC  374

Qy    410 AGTTGATGCCATGGCTGACTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT        469
Db    375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy    470 CTACACTGTCAACCTCTCACAGCAGCTTCATCGTGGAGTCATCAAGCAAGGAT        529
Db    435 CTACACAGTCAACCTCTCACAGCAGTCTCCATCGCCTTCATCAGTGTCGACCGCTA     494

Qy    530 CCTCGCCATTGTCCAGCACCAACAGTCAAAGGCCAAGGAAACTGCTCGGCTGAAAAGGC  589
Db    495 CCTGGCCATCGTCCACCGCCCACCAACAGTCAGGAGGCCAAGGAGCTGTTGGCTGAAAAGGT 554

Qy    590 AGTCTATGTGGGCGTCTGGATCCTTGCCTATGCCTTCCTGGGGCCAAGTCTGTGCCC    649
Db    555 GGTCTATGTTGGCGTCTGGATCCTTGCCTATGCCTTCCTGGGGCCAAGTCTCATCTTTGC 614

Qy    650 CGACGTCAGCAGGGGGACATCAGTCAGTCAGGGGGATGACAGGTACATCGTGACCGCCTTA 709
Db    615 CAACG------------TCAGTGAGGCAGATGACAGATATCTGTGACCGCTCTA       659

Qy    710 CCCCGATAGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCT   769
Db    660 CCCCAATGACTTGTGGGTGGTTGTGTTTCAGTTTCAGTCAGCACATCATGTGGCCTTATCCT 719

Qy    770 GCCCGGCATCGTCATCCTCCGTCTACTGCATATTGCATTATCATCTCCAAGCTGTCACACTCCAA 829
Db    720 GCCTGGTATTGTCATCCTCCGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAA 779

Qy    830 GGGCCACCAGAGCGCAAGGCCCTCAAGACGCAGTCATCCTCATCCTTCTTGC        889
Db    780 GGGCCACCAGAGCGCAAGGCCCTCAAGACACCACAGTGCAACAGCCACCACATCCTTCTTCGC 839

Qy    890 CTGCTGGCTGCCATATTATGTGGGATCAGCCTCCTTCATCCTTTCATCCTTTGGGAGTCAT 949
Db    840 CTGTTGGCTGCCTTACTACATTGGGATCAGCCTCCTTCATCCTCATCCTCTGGGAAATCAT 899

Qy    950 CAAGCAAGGAATGTAGCTTCGAGAGTCATTGTGCACAAGTGGGATCCATCAGAGGGGCCT  1009
Db    900 CAAGCAAGGGTGAGGTTTGAGTAAGACACTGTGCACAAGTGGGATTTCATCACCGAGGGGCCT 959

Qy   1010 CGCTTCTTCCACTGTTGCCTGAACCCCATCCTCTTTCATCCTTTTGGGGCCAAGTTCAA   1069
Db    960 AGTTTCTTCCACTGTTGCTTGTGTCTGAACCCCATCCTCTGTGCTTTCCTTGGAGCCAAATTAA 1019

Qy   1070 AAGCTCTGCCCAGCATGCACTCAACTCCATGAGCAGAGGCTCAAGATCCTTTC       1129
Db   1020 AACCTCTGCCCACCACCACCTCACCTCTGTGAGCAGAGGGTCCAGCTCCAAGATCCTCTC 1079


Qy    481 ACCTCTACAGCGCGGTTCCATCCTGCCTTCATCAGCCTGACCTCGGTACCTCGCCATTG  540
Db   2648 ACCTCTACAGCAGCGGTTCCATCCTGCCTTCATCAGCCTGACGTGCTCCAGTAGCCGCCATTG 2707

Qy    541 TCCACGGCCACCACAGTCAAAGGCCAAGGAAACTGCTCGGCTGGCTGAAAAGGCAGTCTATGTGG 600
Db   2708 TCCACGGCCACCACAGTCAAAGGCCAAGGAAACTGCTGGCTGGCTGAAAAGGCAGTCTATGGG 2767

Qy    601 GCGTCTGGATCCAGCCCTCCTGACTATACCGTGACTCTTTGCCGACGTCAGCC        660
Db   2768 GCGTCTGGATCCTCCCTCCTGACTATACCGTGACTCTTTGCCGACGTCAGCC         2827

Qy    661 AGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCGCCTTTACCCCGATAGCC    720
Db   2828 AGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCGCCTTTACCCCGATAGCC   2887

Qy    721 TGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCG   780
Db   2888 TGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCG   2947

Qy    781 TCATCCTCCTGTACTGACAGTCATCATCATCTAGCTGCACACTCCAGGGGCCACCAGA   840
Db   2948 TCATCCTCCCTGTTACTGACAGTCATCATCATCTCCAAGCTGTCCACACTGTCCAGGGCCACCAGA 3007

Qy    841 AGGCCAAGGCCCTCAAGACGCAGTCATCCTCATCCTCATTCTTTGCCTGCTGGCTGC    900
Db   3008 AGGCCAAGGCCCTCAAGACGCAGTCATCCTCATCCTCAGTTCTTTGCCTGCTGGCTGC   3067

Qy    901 CATATTATGTGGGGATCATCAGCCTCCTTCATCCTTTTTGGGAGTCATCAAGCAAGGAT  960
Db   3068 CATATTATGTGGGGATCAGCCATCCACCATCCTTCTTTTGGGGGTCATCAAGCAAGGAT  3127

Qy    961 GTGACTTCGAGAGCAGCATTGTGCACAAGTGCACAAGTCGCCCTCCGCCTTCTTCC     1020
Db   3128 GTGACTTCGAGAGCAGCATCGTGCACAAGTGTGCACAAGTCAGGCCCCTCGCCTTCTTCC 3187

Qy   1021 ACTGTTGCCTGAACCCCATCCTCTTATGCCTTCCTGGGGCCAAGTCTGCCC         1080
Db   3188 ACTGTTGCCTGAACCCCATCCTCTCTATGCCTTCCTGCGGGCCAAGTCTCAAAAGCTCTGCCC 3247

Qy   1081 AGCATGACACTCAACTCAACTCTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTCCAGTTTCAAGATCCTGGA 1140
Db   3248 AGCATGACACTCAACTCAACTCTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTGGA    3307

Qy   1141 GGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAA   1199
Db   3308 GGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAA   3366
```

RESULT 14
AR015970

Query Match   51.9%;   Score 974.4;   DB 6;   Length 1737;
Best Local Similarity   76.1%;   Pred. No. 5.7e-192;

Query Match   51.9%;   Score 974.4;   DB 6;   Length 1737;
Best Local Similarity   76.1%;   Pred. No. 5.7e-192;
Matches 1391;   Conservative   0;   Mismatches 301;   Indels 136;   Gaps   10;

```
Qy     50 GCAGGTAGCAGTGACCCTGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC 109
          |||||||||| |||| || |||   |||||||||||||||||||| |||
Db     20 GCGCGGCGCAAGTGACGCCGAGGGCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy    110 GAGTGTTGCCATGGAACGATCAGTGAGTATATACACTCTGATAACTACTCTGAGA 169
          ||||| | |  ||||||||||||||||| |||||||||||| ||
Db     80 CAGCCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACCGAGGA 134

Qy    170 AGTGGGCTTGGAGACTATGACTCCAACAAGGACCCTGCTTCCGGGATGAAACGTCCA 229
          ||||| |||||||||||||||||||| ||||||| |||||||||||||||
Db    135 AATGGGCTCAGGGGACTATGACTCCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy    230 TTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTGACTGGCATAGTCGG 289
          |||||||  || ||||||||||||||||  ||||| ||||| || |||||||| |||
Db    195 TTTCAATAAAATCTTCCTGCCACCATCTACTTCTTCTTCTTAACTGGCATTGTGGG 254

Qy    290 CAATGGATTGGTGATCCTGGTCATGGGTTACCAGGAGCTAGGGAGCATGACGGACAA 349
          |||||||||||||||||||||||||||||||||||||||||| |||||| |||||||
Db    255 CAATGGATTGGTCATCCTGGTCATGGGTTACCAGGAGAAGCATGACGGACAA 314

Qy    350 GTACCGGCTGCACCTGTCAGTGGTCGACCTCCCTTGTCATCACACCCCTTCTGGGC 409
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 GTACAGGCTGCACCTGTCAGTGGCCGACCTCCTTGTCATCACGCTTCCTGTTCTGGGC 374

Qy    410 AGTTGATGCCATGGCCGACTTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 AGTTGATGCCGTGGCCTGGCAAACTCGTACTTTGGGAACTTCCTATGCAGGCAGTCCATGTCAT 434

Qy    470 CTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTA 529
          |||||||||||||||||||||||||| ||||||||||||||||||||||||| |||||
Db    435 CTACACAGTCAACCTCTACAGCAGCGTGTCTCATCCTGGCCTTCATCAGTCTGGACCGGCTA 494

Qy    530 CCTGGCCATTGTCCACGCCACCAAGTGCTGGCTGAAAAAGGC 589
          |||||||||||||||||||||||||||||||||||||||
Db    495 CCTGGCCATCGTCCACGCCACCAAGAACCAGTCATGCTGGTTGCTGAAAAAGGT 554

Qy    590 AGTCATGTGGGCGTCTGGATCCCAGGCCCCTCCTGACTAGATTGACTTCAGTCCTACAAAT 649
          ||||||||||| |||||||||||| || |||||||||||||||||||
Db    555 GGTCTATGTTGCTGCGTCTGGATCCCTGCCCTCCTGCGACTATTCCCGACTTCATCTTTGC 614

Qy    650 CGACGTCAGCCAGGGGACATCAGTCAGGGGATGAGACAGTGACATCGTGTCGCCTTTA 709
          |||    |||                      ||||||||||||||||||||| |||||||
Db    615 CAACG-----------------------TCAGTGAGGCAGATGACAGATAATCTGTGACCGCTTCTA 659

Qy    710 CCCCGATAGCCTGTGATGGTGGTCTGTTCAATTCGACCATCATCATCTCTAAGCGTCACACTCCAA 769
          |||| |||||||||||||||||||||||||||||||||||||||||
Db    660 CCCCAATGACTTGTGGGTGGTTGTGTTCCAGTTCAGCAATAATCATCATCATGGTTGCCTTATCCT 719

Qy    770 GCCCGGCATCGTCATCCTCTCCGTTACTGCTCATCATCTCTAAGCGTCGTCACACTCCAA 829
          ||| || |||| |||||||||||||||||||||||||||||||||||||||
Db    720 GCCTGGTATTGTCATCCTGTCATCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAA 779

Qy    830 GGGCCACCAAGAGCGCAAGGCCCTCAAGAGACAGCAGTCATCCTCATCCTCAGTTTCTTTGC 889
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    780 GGGCCACCAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTCGGCTTTCTTCTTCGC 839

Qy    890 CTGCTGGCTGCCATATTATGTGGGGATCAGCGATCAGCATGACTCCTTCATCCTTTTGGGAGTCAT 949
          |||||||||| |||||||||||||||| ||||||||||||||||||||
Db    840 CTGTTGGCTGCCTTACTACACATTGGGATCAGCATCAGCGATCAGGATGACTCCTTCATCCTCTGGAAATCAT 899

Qy    950 CAAGCAAGGATGTGACTTGACAAGTGACATTGTGCACAAGTGGATCTCCATCACAGAGGCCCT 1009
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    900 CAAGCAAGGGTGTGAGTTTGAGACAACTGTGCACACTGTGCACAAGTGGATTCCATCACCGAGGCCCT 959

Qy   1010 CGCCTTCTTCCACTGTGCCTGAACCCCATCCTCTCTATGCCTTCCTCGGGGCCAAGTTCAA 1069
          || ||||||||||||||||||| |||||||||||||||||||||||||||
Db    960 AGTTGGCTGCCTTCTTCCACTGTGCTGTGAACCCCACTCTTCATCCTCTATGCCTTCCTATGGGGAGCCAAATTTAA 1019

Qy   1130 CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGAGTCAGAATCCTCCAGTTTCA 1189
          |||||||||||||||||||||||| |  |||||||||||||| |||||||||||
Db   1080 CAAGGGAAAGCGCGGGTGGACATTCATCATCTGTTTCCACTGAGTCTGAGTCTTCAAGTTTCA 1139

Qy   1190 CTCCAGCTAACCCTATGCAAAGACTTATATATAATATATATATATGATAAAGAACTT 1249
          |||||||   |||| ||||||||||| ||||  ||||||||||| ||||||||||||
Db   1140 CTCCAGCTAA------CACAGATGTAAAAGACTTTTTTTTTTTATACGATAAATAACTTTT 1191

Qy   1250 TTTTATGTTACACATTTTCAGATATAAGAGACTGACCAGTCTGTACAGTTTTTTTTTT 1309
          |||||||||||| ||||||||||||||||||||||||| ||||||||||||||||
Db   1192 TTTAAGTTACACATTTTTCAGATATAAAGAGACTGACCAATATTGTACAGTTTTTATTGC 1251

Qy   1310 TTTTTAATTGACTGTGTGGGAGTTTATGTTCCTCCAGTTTTGTGACGTTTGACTTAATTT 1369
          ||||||||||||| |||  |||  ||||  |||||||||||||||| |||||  |
Db   1252 TTGTTGGGATTTTTGT-----CTTGTGTTTCTTTAGTTTTTGTGAAGTTTTAATTGACTTA 1305

Qy   1370 ATATAAATAATATGTTTTGTTGTTCATGTGAATGAGCGGTCTAGGCAGGACCTGTGGCC 1429
          ||||||      ||||  ||||||||||||||||||||| |||| |||||||||||||||
Db   1306 TTTATA---TAAATTTTTTTGTTGTTCATATTGATGTGTCTCAGGCAGGACCTGTGGCC 1361

Qy   1430 AAGTTCTTAGTAGCTGTTATTCTGTGTGTAGGACTGTAGAAACTGTAGAGGAAGAAACTGA 1489
          ||||||||||||||||||||||||||||||||||||||||      ||||||||||
Db   1362 AAGTTCTTAGTTGCTGTCTCTGCCTGTGTGCTGGACTGTGAGA------AAGGGACTGA 1413

Qy   1490 ACATTCCAGAATGTGTGTAAATTGATGAAAGTAGCCGTGATTGTTGCTGCA 1549
          ||||||||||||||| |||||||||||||||| |||||| |||||||||||
Db   1414 ACATTCCAGAGCGTGTAGTGAGTGAATCACGTAAAGCTAAGAAATGATCCCCAGCTGTTTATGCA 1473

Qy   1550 TAATCTCTTCATTCCGAGGAGCACCCCACCCCCACCCCCCATTCTTAAAT 1609
          ||||| |||||||||| |
Db   1474 TAGATAATCTCTCCA-------------------------------------- 1488

Qy   1610 TGTTTGGTTATGCTGTGTGATGGTTTGTTTTTTTTTGTGTTGTTTTT 1669
          ||||||||||||||||||||| ||   ||||||||
Db   1489 ----------------------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGAT 1525

Qy   1670 TTTTTCGTAAAGATGGCACTTAAAAGCCAAAGGCTGAAAATGCTGGGGT 1729
          |||||||   |||||||||||||| ||||||||||
Db   1526 TTTGCTGTAGAAGATGGCACTTATAACCAAAGCCCAAAGTGT-ATAGAAATGCTG--- 1581

Qy   1730 TTTTTTGTTTTGTTTTTTCAGTTTCAAGAGAGTTGACTTCAGTCCTACAAAT 1789
          ||||||||  |||||||||||||||||||||| ||||||||||||||
Db   1582 ------------TTTTTCAGTTTTCAGGAGTGGGTTGATTTCAGCCACACTAC-AGT 1623

Qy   1790 GTACAGTCTTGTATTACATTGTTAATAAAGTCAATGATAAAACTTAAAAAAAAAAAA 1849
          |||||||||||||||||||||||||||| ||||| |||||  || ||||||||||||
Db   1624 GTACAGTCTTGTATTAAGTTGTTAATAAAGTACACATGTTAAACTTAAAAAAAAAAAAAA 1683

Qy   1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
          |||||||||||||||||||||
Db   1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

RESULT 15
AR060748
LOCUS   AR060748   1737 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION   Sequence 3 from patent US 5840856.
ACCESSION   AR060748
VERSION   AR060748.1   GI:5987198
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1737)
  AUTHORS   Chuntharapai,A., Lee,J., Hebert,C. and Kim,K.Jin.
  TITLE   Antibodies to a human PF4 superfamily receptor
  JOURNAL   Patent: US 5840856-A 3 24-NOV-1998;
FEATURES   Location/Qualifiers
     source   1..1737
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

```
Qy  1070  AAGCTCTGCCCAGCATGCACTCAACTCCATGAGCAGGAGGCTCCAGCCTCAAGATCCTTTC  1129
                ||||||||||| |||| | || | ||||||||| ||||||||||||||||||||| ||
Db  1020  AACCTCCTGCCCCAGCACGCACTCACCTCGTGTGCAGAGGGTCCAGCCTCAAGATCCTCTC  1079

Qy  1130  CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTTCA  1189
                |||||||||| |||||| ||| || | ||| |  | || |||  |||||||||||||
Db  1080  CAAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTTCTTCAAGTTTTCA  1139

Qy  1190  CTCCAGTAACCCTTATGCAAAGACTTATATATATATATATATATGATAAAGAACTT  1249
                |||||||||                 | |
Db  1140  CTCCAGCTAA-------CACAGATGTAAAAGACTTTTTTTTATACGATAAATAACTTTT  1191

Qy  1250  TTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTT  1309
                |||||  || | |||| ||||||||||||||  || |||| | ||||||  ||||||||
Db  1192  TTTTAAGTTACACATTTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTTTTATTGC  1251

Qy  1310  TTTTTAATTGACTGTTGGGAGTTATGTTCCTCTAGTTTTGTGAGGTTGACTTAATTT  1369
                |||    |||| |  |||  |||||| || || |||||| |||||  | ||| |
Db  1252  TTTTGGATTTTTGT-----CTTGTGTTTCTTTAGTTTTGTGAAGTTTAATTGACTTA  1305

Qy  1370  ATATAAATATGTTTTTTGTTCATCGTGTGTAGGAGCGTCTAGGACGACCTGTGGCC  1429
                |||||    ||||||||||  |||||   |||||  ||||||| ||||||||||||
Db  1306  TTTATA---TAAATTTTTTTGTTTCATATTGATGTGTCTAGGAGGACCTGTGGCC  1361

Qy  1430  AAGTTCTTAGTAGCTGTTTATCGTGTGTAGGAGACTGTAGAACTGTAGAGGAAGAAACTGA  1489
                |||||||||| ||| ||||| || || || ||  | ||||        |||| ||||||
Db  1362  AAGTTCTTAGTTGCTGTGATGTCTCGTGTAGGACTGTAGAA-------AAGGGAACTGA  1413

Qy  1490  ACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCGTGTTGCTGCA  1549
                ||||||||||  || | || |  |||| | |||| |||| |||||||||| || ||||||
Db  1414  ACATTCCGAGAGCGTGTAGTGAGTGAATCAGCTAGAAAATGATCCCCAGCTGTTTTATGCA  1473

Qy  1550  TAATCTCTTCATTCCGAGGAGCACCCCACCCCCACCCCACCCCATTCTTAAAT  1609
                ||
Db  1474  TAGATAATCTCTCCA--------------------------------------  1488

Qy  1610  TGTTTGGTTATGCTGTGTGATGGTTGTTGGTTTTTTTTGTTGTTGTTGTTTTTT  1669
                  ||    ||||| |  | ||||||||| |
Db  1489  --------------TTCCCGTGAACGTTTTCCTGTTCTTAAGACGTGAT  1525

Qy  1670  TTTTCTGTAAAAGATGGCACTTAAAACCAAAGCCTGAAATGGTGGTAGAAATGCTGGGGT  1729
                ||||| || ||| ||||| ||| |||| |||||||| | |||  |||||||||| |||
Db  1526  TTTGCTGTAGAAGATGGCACTTATAACCAAAGCCCAAGTGGT-ATAGAAATGCTGG---  1581

Qy  1730  TTTTTTTGTTTGTTGTTTTTTCAGTTTTCAAGAGTAGATTGACTTCCAGTCCCTACAAAT  1789
                             |||||||||  ||| |||| | |||||
Db  1582  --------------TTTTTCAGTTTTCAGGAGTGGGTTGATTTCAGCACCTAC-AGT  1623

Qy  1790  GTACAGTCGTCTTGTATTACATTGTTAATAAAAGTCAATGATAAACTTAAAAAAAAAAAA  1849
                ||||||| |||| |||||||||| ||||||||||| |
Db  1624  GTACAGTCTTGTTATTAAGTTGTTAATAAAAGTACATGTTAAACTTAAAAAAAAAAAAAAA  1683

Qy  1850  AAAAAAAAAAAAAAAAAAAAAAAAA  1877
                |||||||||||||||||||||||||
Db  1684  AAAAAAAAAAAAAAAAAAAAAAAA  1711
```

Search completed: May 17, 2004, 12:10:40
Job time : 7607.63 secs

```
                GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      May 17, 2004, 03:46:58 ; Search time 804.917 Seconds
                                    (without alignments)
                            9906.453 Million cell updates/sec

Title:       US-09-367-052-1
Perfect score: 1877
Sequence:    1 ccatcctaatacgactcact.......aaaaaaaaaaaaaaaaaa 1877

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched:    3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:      6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq.29Jan04:*
             1: geneseqn19809:*
             2: geneseqn19990s:*
             3: geneseqn20000s:*
             4: geneseqn20001as:*
             5: geneseqn20001bs:*
             6: geneseqn20002s:*
             7: geneseqn20003as:*
             8: geneseqn20003bs:*
             9: geneseqn20003cs:*
            10: geneseqn20004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                        SUMMARIES
```

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1875.4	99.9	1877	2	AAV46370	Aav46370 Nucleic a
2	1758	93.7	1758	2	AAZ27611	Aaz27611 Mouse CXC
3	974.4	51.9	1737	2	AAQ80521	Aaq80521 Human mon
4	974.4	51.9	1737	2	AAQ99007	Aaq99007 Chemokine
5	969.6	51.7	1737	2	AAQ29506	Aaq29506 New plate
6	952.8	50.8	1679	4	AAD12801	Aad12801 Human neu
7	952.8	50.8	1679	4	AAI65467	Aai65467 Nucleotid
8	952.8	50.8	1679	6	ABS53992	Abs53992 DNA encod
9	952.8	50.8	1679	6	ABV78039	Abv78039 Hypoxia-r
10	952.8	50.8	1679	6	ABX08779	Abx08779 Angiogene
11	952.8	50.8	1679	7	ABX74454	Abx74454 Human cDN
12	952.8	50.8	1679	7	ABZ68886	Abz68886 Nucleotid
13	952.8	50.8	1679	7	ABZ42642	Abz42642 Human CXC
14	952.8	50.8	1679	9	ADC98645	Adc98645 Human CXC
15	952.8	50.8	1711	6	ABZ35348	Abz35348 Human gen
16	952.8	50.8	1711	8	ADB47320	Adb47320 Human cDN
17	951.8	50.7	1664	3	AAA61656	Aaa61656 Human CXC
18	951.8	50.7	1664	8	AAZ40014	Aaz40014 CXCR4 cod
19	950.8	50.7	1679	6	ABS53750	Abs53750 DNA encod
20	944.8	50.3	1670	6	ABK83803	Abk83803 Human cDN
21	944.8	50.3	1670	6	ABN95645	Abn95645 Gene #214
22	944.8	50.3	1670	6	ACC46765	Acc46765 Human COP
23	939.8	50.1	1944	2	AAX15882	Aax15882 cDNA enco

24	936.8	49.9	1588	2	AA227610	Aaz27610 Human CXC
25	934	49.8	5161	3	AAA34774	Aaa34774 Human ade
26	934	49.8	5161	3	AAF20896	Aaf20896 Human CCR
27	934	49.8	5161	6	AAH99951	Aah99951 CXCR4 gen
28	934	49.8	5161	7	ABZ96590	Abz96590 Human CCR
29	934	49.8	5161	7	ACA64841	Aca64841 Human CXC
30	933.6	49.7	1582	6	ABS64793	Abs64793 CXCR4 rec
31	932.4	49.7	8747	9	ADD67548	Add67548 Human Ly1
32	905.6	48.2	1050	9	ADB58421	Adb58421 Toxicity-
33	905.6	48.2	1050	9	ADB53005	Adb53005 Primary r
34	859	45.8	1225	6	ABZ35630	Abz35630 Human gen
35	859	45.8	1225	7	ACA56637	Aca56637 Human sig
36	847	45.1	1224	6	ABZ35563	Abz35563 Human gen
37	834.6	44.5	1317	6	AAV18357	Aav18357 Human RM3
38	834.6	44.5	1317	3	AAA91726	Aaa91726 Human 7TM
39	834.6	44.5	1317	6	ABK54255	Abk54255 Human 7 t
40	833	44.4	1317	6	AAQ66179	Aaq66179 Seven tra
41	809.6	43.1	1059	3	AAZ38553	Aaz38553 Human CXC
42	809.6	43.1	1059	6	ABQ74938	Abq74938 Human CXC
43	809.6	43.1	1059	6	AAH99952	Aah99952 CXCR4 enc
44	808	43.0	1059	7	ACC72707	Acc72707 Human can
45	808	43.0	1102	7	ACA64750	Aca64750 Chemokine

                                ALIGNMENTS

```
RESULT 1
AAV46370
ID   AAV46370 standard; cDNA to mRNA; 1877 BP.
XX
AC   AAV46370;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Nucleic acid encoding a murine CXC chemokine receptor.
XX
KW   Mouse; CXC chemokine receptor; pre-B cell line DW34;
KW   CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1; HIV infection;
KW   screening; inhibitor; AIDS; ds.
XX
OS   Mus sp.
XX
FH            Key             Location/Qualifiers
FT            CDS             120..1199
FT                            /*tag= a
XX
PN   WO9835035-A1.
XX
PD   13-AUG-1998.
XX
PF   07-FEB-1997;  97WO-JP000299.
XX
PR   07-FEB-1997;  97WO-JP000299.
XX
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;
PI   Nakajima T, Yoshie O;
XX
DR   WPI; 1998-447232/38.
DR   P-PSDB; AAW64778.
XX
PT   Mouse CXC chemokine receptor binding to PBSF/SDF-1 pre-B cell stimulatory
PT   factor - is useful for screening of potential HIV infection and AIDS
PT   inhibitors.
XX
PS   Claim 3; Page 39-42; 76pp; Japanese.
XX
CC   The present sequence encodes a murine CXC chemokine receptor which binds
CC   to the mouse CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1. The
CC   nucleic acid is isolated from mouse pre-B cell line DW34. The receptor
CC   and cells expressing it can be used in the study and mapping of the
```

CC   mechanism of HIV infection and in screening of potential inhibitors of
CC   HIV infection and the development of AIDS
XX
SQ   Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T; 0 U; 0 Other;

Query Match   99.9%; Score 1875.4; DB 2; Length 1877;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1876; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

```
Qy      1 CCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGTGCAGGTAGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGTGCAGGTAGCAG 60

Qy     61 TGACCCCTGAGGCGTTGGTGTCTCCGGTAACCACCACGGCGTGTAGAGCGAGTGTTGCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGACCCCTGAGGCGTTGGTGTCTCCGGTAACCACCACGGCGTGTAGAGCGAGTGTTGCCA 120

Qy    121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG 180

Qy    181 GAGACTATGACTCCAACAAGGAACCCGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAGACTATGACTCCAACAAGGAACCCGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 240

Qy    241 TCTTCCTGCCCCACCATCTACTTCATCATCTTCTGACTGCATAGTCGGCAATGGATTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCTTCCTGCCCCACCATCTACTTCATCATCTTCTGACTGCATAGTCGGCAATGGATTGG 300

Qy    301 TGATCCTGGTCATGGGTACTTCCAGAGAAGAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGATCCTGGTCATGGGTACTTCCAGAGAAGAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 360

Qy    361 ACCTGTCAGTGGCTGACCCTCTTGTCATCACACACTCCTTGGGCAGTTGATGCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ACCTGTCAGTGGCTGACCCTCTTGTCATCACACACTCCTTGGGCAGTTGATGCCA 420

Qy    421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480

Qy    481 ACCTCTACAGACAGCGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCCATTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCTCTACAGACAGCGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCCATTG 540

Qy    541 TCCACGCCACCAACAGTGCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCCACGCCACCAACAGTGCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600

Qy    601 GCGTCTGATCCCAGCCCTCCCGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCGTCTGATCCCAGCCCTCCCGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 660

Qy    661 AGGGGGACATCAGTAGTGAGGGATGACAGTACATCTGTGACCGTCCTTACCCCGATAGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AGGGGGACATCAGTAGTGAGGGATGACAGTACATCTGTGACCGTCCTTACCCCGATAGCC 720

Qy    721 TGTGGATGTGGTGTTCAATTCCAGCATATAATGGTGGTCTCATCCTGCCCGGCATCG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGTGGATGTGGTGTTCAATTCCAGCATATAATGGTGGTCTCATCCTGCCCGGCATCG 780

Qy    781 TCATCCTCCTCGTTACTGCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TCATCCTCCTCGTTACTGCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGA 840

Qy    841 AGCGCAAGGCCCTCAAGACGCAGACGTCATCCTCATCCTCCAGCTTTCTTTGCCTGGCTGC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGCGCAAGGCCCTCAAGACGCAGACGTCATCCTCATCCTCCAGCTTTCTTTGCCTGGCTGC 900

Qy    901 CATATTATGTGGGGGATCAGCATCAGCATCAGCGAGTCATCAAGCAAGGAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CATATTATGTGGGGGATCAGCATCAGCATCAGCGAGTCATCAAGCAAGGAT 960

Qy    961 GTGACTTCGGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTGACTTCGGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCC 1020

Qy   1021 ACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCC 1080

Qy   1081 AGCAATGCACTCAACTCCATGAGCAGCAGAGGCTCAGCCTCAAGATCCTTCCAAAGGAAAGC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGCAATGCACTCAACTCCATGAGCAGCAGAGGCTCAGCCTCAAGATCCTTCCAAAGGAAAGC 1140

Qy   1141 GGGGTGGACACTCTTCCGTCTCCACGAGTCAGAATCCTCCAGTTTCACTCCAGCTAAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCGGTGGACACTCTTCCGTCTCCACGAGTCAGAATCCTCCAGTTTCACTCCAGCTAAC 1200

Qy   1201 CCTTATGCAAAGACTTATATATAATATATATATATATGATAAAGAACTTTTTATGTTAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCTTATGCAAAGACTTATATATAATATATATATATATGATAAAGAACTTTTTATGTTAC 1260

Qy   1261 ACATTTCCAGATATAAGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTAATTGA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ACATTTCCAGATATAAGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTAATTGA 1320

Qy   1321 CTGTGTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGACTTGACTTTATATAAATATT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGTGTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGACTTGACTTTATATAAATATT 1380

Qy   1381 GTTTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGCCAAGTTCTTAGT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGCCAAGTTCTTAGT 1440

Qy   1441 AGCTGTTTATCTGTGTGTAGGACTGTAGAGAGAAGAACTGAACATTCCAGAA 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AGCTGTTTATCTGTGTGTAGGACTGTAGAGAGAAGAACTGAACATTCCAGAA 1500

Qy   1501 TGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTGCATAAATCTCTTCA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTGCATAAATCTCTTCA 1560

Qy   1561 TCCGAGGAGCACCCCACCCCCACCCCCATTCTTAAATTGTTGGTTAT 1620
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TCCGAGGAGCACCCCACCCCCACCCCCATTCTTAAATTGTTGGTTAT 1620

Qy   1621 GCTGTGTGATGGTTTGTTGGTTTTTTGTTGTTTGTTTCTGTAAA 1680
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCTGTGTGATGGTTTGTTGGTTTTTTGTTGTTTGTTTCTGTAAA 1680

Qy   1681 AGATGCACTTAAAACCAAAGCCTGAAATGTGGTAGAAATGCTGGGGTTTTTTGTT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 AGATGCACTTAAAACCAAAGCCTGAAATGTGGTAGAAATGCTGGGGTTTTTTGTT 1740

Qy   1741 GTTTTGTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GTTTTGTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTG 1800

Qy   1801 TATTACATTGTTAATAAAAGTCAATGATAAACTTAAAAAAAAAAAAAAAAAAAAAA 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TATTACATTGTTAATAAAAGTCAATGATAAACTTAAAAAAAAAAAAAAAAAAAAAA 1860

Qy   1861 AAAAAAAAAAAAAA 1877
          ||||||||||||||
Db   1861 AAAAAAAAAAAAAA 1877
```

RESULT 2
AAZ27611
ID   AAZ27611 standard; DNA; 1758 BP.
XX
AC   AAZ27611;
XX
DT   16-DEC-1999 (first entry)
XX

DE   Mouse CXCR4 coding sequence.
XX
KW   CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;
KW   tissue repairing agent; vascularisation; ss.
XX
OS   Mus sp.
XX
PN   WO9948528-A1.
XX
PD   30-SEP-1999.
XX
PF   23-MAR-1999;   99WO-JP001448.
XX
PR   24-MAR-1998;   98JP-00095448.
XX
PA   (CHUS) CHUGAI SEIYAKU KK.
PA   (KISH/) KISHIMOTO T.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K;
XX
DR   WPI; 1999-591042/50.
DR   P-PSDB; AAY39994.
XX
PT   CXCR4-potentiating agents and methods useful for inhibiting
PT   neovascularization, and treating solid cancers.
XX
PS   Disclosure; Page 50-51; 63pp; Japanese.
XX
CC   This sequence encodes the mouse CXCR4 protein. The invention relates to
CC   remedies inhibiting neovascularisation, remedies for solid cancer,
CC   remedies for diseases pathologically caused by neovascularisation and
CC   tissue repairing agents containing the active ingredient a substance
CC   capable of potentiating CXCR4. Based on a finding that vascularisation is
CC   inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC   remedies inhibiting vascularisation which contain as the active
CC   ingredient a substance capable of potentiating CXCR4, remedies for solid
CC   cancer, remedies for diseases pathologically caused by neovascularisation
CC   and tissue repairing agents containing as the active ingredient a
CC   substance capable of potentiating CXCR4. It is also possible to establish
CC   methods for treatment with the use of these remedies
XX
SQ   Sequence 1758 BP; 447 A; 414 C; 369 G; 528 T; 0 U; 0 Other;

Query Match   93.7%;   Score 1758;   DB 2;   Length 1758;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1758;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    120 ATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCT 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCT 60

Qy    180 GGAGACTCAACTTCCAACAAGGAACCCTGCTTCCGGGATGAAAAACGTTCCATTCAATAGG 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGAGACTCAACTTCCAACAAGGAACCCTGCTTCCGGGATGAAAAACGTCCATTCAATAGG 120

Qy    240 ATCTTCCTGCCCACCATCACTTCCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTG 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATCTTCCTGCCCACCATCACTTCCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTG 180

Qy    300 GTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGCAGTACCGCTG 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGCAGTACCGGCTG 240

Qy    360 CACCTGTCAGTGCTGCTGACCTGACGACACTCCCCTTCTGGCCAGTTGATGCC 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CACCTGTCAGTGCTGCTGACCTGACGACACTCCCCTTCTGGCCAGTTGATGCC 300

Qy    420 ATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCGTGTCCATATCATCATCTTC 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCGTGTCCATATCATCACACTGTC 360

Qy    480 AACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGACGTACCTGCCATT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGACGTACCTGCCATT 420

Qy    540 GTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTG 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTG 480

Qy    600 GGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACTCAGC 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGCGTCTGGATCCCAGCCCCTCCTGACTATACCTGACTTCATCTTTGCCGACTCAGC 540

Qy    660 CAGGGGGACATCAGTCAGGGGATGACAGCAGGTACATCTGTGACCGCCTTTACCCGATAGC 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGGGGGACATCAGTCAGGGGATGACAGCAGGTACATCTGTGACCGCCTTTACCCGATAGC 600

Qy    720 CTGTGGATGGTGTCTTCAATTCCAGCATATAATGTGGTCATCCTGCCCGGCATC 779
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGTGGATGGTGGTTTCAATTCCAGCATATAATGTGGTCTCATCCTGCCCGGCATC 660

Qy    780 GTCATCCTCTCCTGTTACTGTCATCATCATCTCAAGCTGTCACACTCCAAGGGCCACCAG 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTCATCCTCTCCTGTTACTGCATCATCATCTCAAGCTGTCACACTCCAAGGGCCACCAG 720

Qy    840 AAGGCAAGGCCCTCAAGACGACAGTCATCCTCCTTCATCCTCATCCTCTAGCTTTCTTTGCCTGCTG 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AAGGCCAAGGCCCTCAAGACGACAGTCATCCTCCTTCATCCTCTAGCTTTCTTTGCCTGCTG 780

Qy    900 CCATATTATGTGGGGATCAGCATGCATCGACTCCTTCCTCCATCCTCCTTTTGGGAGTCATCAAGCAAGGA 959
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CCATATTATGTGGGGATCAGCATGCATCGACTCCTTCCTCCATCCTCCTTTTGGGAGTCATCAAGCAAGGA 840

Qy    960 TGTGACTTGGAGAGCATTGTGCACAAGTGACTTCCATCACAGAGGCCCGCCGCTTCTTC 1019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGTGACTTGGAGAGCATTGTGCACAAGTGGATCTTCCATCACAGAGGCCCCGCCTTCTTC 900

Qy   1020 CACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTCGGGGGCCAAGTTCAAAAGTCTGCC 1079
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTCGGGGGCCAAGTTCAAAAGTCTGCC 960

Qy   1080 CAGCATGACTCACTCCATGAGCAGAGGCTCAGCCTCCAAGATCCTTCCAAAGGAAAG 1139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CAGCATGACTCACTCCATGAGCAGAGGCTCAGCCTCCAAGATCCTTCCAAAGGAAAG 1020

Qy   1140 CGGGGTGGACACTCCTCCGTCCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAA 1199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CGGGGTGGACACTCCTCCGTCCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAA 1080

Qy   1200 CCCTTATGCAAAGACTTATATAATAATAATATATATATATATATGAGAACTTTTTATGTTA 1259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCCTTATGCAAAGACTTATATAATAATAATATATATATATATATGAGAACTTTTTATGTTA 1140

Qy   1260 CACATTTTCCAGATATAAAGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTAATTG 1319
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CACATTTTCCAGATATAAAGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTAATTG 1200

Qy   1320 ACTGTTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGACTTAATTATATAAAATAT 1379
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACTGTTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGACTTAATTATATAAAATAT 1260

Qy   1380 TGTTTTTTGTTGTTCATGTGAATGAATGAGGCTCAGGCAGGACCTGTGGCCAAGTTCTTAG 1439
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGTTTTTTGTTGTTCATGTGAATGAATGAGGCTCAGGCAGGACCTGTGGCCAAGTTCTTAG 1320

Qy   1440 TAGCTGTTTATCTGTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGA 1499
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TAGCTGTTTATCTGTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGA 1380

Qy   1500 ATGTGTGGTAAATTGGATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCGATAATCCTTC 1559
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ATGTGTGGTAAATTGGATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCGATAATCCTTC 1440

Qy   1560 ATTCCGAGGAGCACCCCACCCCACCCCCACCCCCCATTCTTAAATTGTTTGGTTA 1619
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATTCCGAGGAGCACCCCACCCCACCCCCACCCCCCATTCTTAAATTGTTTGGTTA 1500
```

```
Query Match          51.9%;  Score 974.4;  DB 2;  Length 1737;
Best Local Similarity  76.1%;  Pred. No. 2.9e-178;
Matches 1391;  Mismatches   0;  Conservative   0;  Indels 136;  Gaps  10;

Qy      50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACCGGCTGTGAGAGC 109
           ||||||||||||||||||||||| ||||||||||||   |||||||| ||||||||||||
Db      20 GCGCGGCGCAAAGTGACGCCCAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGAGAAAC  79

Qy     110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
           ||| |||||||||||||||||||||||||     ||||||||||||||||||||||||||
Db      80 CAGCGGTTACCATGGA-----GGGGATCAGTAGTCAGTATATACACTTCAGATAACTACACCGGAGGA 134

Qy     170 AGTGGGGTCTGAGACTATGACTCCAACAACGAACCCGCTTCCGGGATGAAAACGTCCA 229
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     135 AATGGGGCTCAGGGGACTATGACTCCATGGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy     230 TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGG 289
           ||||||||||||| ||||||||||||||||||||||||||||| |||||||||||||||
Db     195 TTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTTCTTAACTGGCATTGTGGG 254

Qy     290 CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAGAGAAGAGTAAGGACGATGACGGACAA 349
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 CAATGGATTGGTCATCCTGGTCATGGGTTACCAGAGAGAAAACTGAGAGAGCATGACGGACAA 314

Qy     350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTGGGC 409
           |||||||||||||||||||||||||||||||||||||| ||||||||| |||||||||||
Db     315 GTACAGGCTGCACCTGTCAGTGCCGACCTCCTCTTGTCATCACGGCTTCCCCTTCTGGGC 374

Qy     410 AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCAT 469
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy     470 CTACACTGTCAACCTCTACAGACAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGTA 529
           ||||||||| ||||| |||||||||||||||||| |||||||||||||||||||||||
Db     435 CTACACAGTCAACCTCTACAGACAGCGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGCTA 494

Qy     530 CCTCGCCATTGTCCACGCCACCACCAACAGTCAAAGGCAAGGACCAAGGAAACTGCTGGCTGAAAAGGC 589
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     495 CCTGGCCATTGTCCACGCCACCACCAACAGTCAGAATGGCCAAGGAAGCAGTTGGCTGAAAAGGT 554

Qy     590 AGTCTATGTGGGGCGTCTTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGC 649
           |||| ||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db     555 GGTCTATGTGGGGCGTCTGGATCTCCTGCCCTCCTGCTGACTAATCCCGACTTCATCTTTGC 614

Qy     650 CGACGTCAGCCAGCAGGGGGACATCAGTCAGGGGATGACAGCAGGTACATCTGTGACCGGCCTTA 709
           |||||||||||||||||||||||||||||        |||||||||||||| |||||||||||
Db     615 CAACG-----------TCAGTGAGGCAGAGCAAGTGGTGTTTCAAGGTCAGCAGTCAGATATATCTGTGACCGGCTCTA 659

Qy     710 CCCCGATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATATAATGGTGGGTCTCATCCT 769
           ||||||||||||||||||||||||||||| |||||||| ||| |||||||||||||||||
Db     660 CCCCAATGACTTGTGGGTGGTGGTTTCCAGTTTCAGCTTCAGCACACAATCATGCATGGGTCTTGGCCTTATCCT 719

Qy     770 GCCCGGCATCGTCATCCTCCTGTTACTGCATCATCATTCTCAAGGTGTCACACTCCAA 829
           |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db     720 GCCTGGTATTGTCATCCTCGTCTATTGCATTTGCATTTATCATCTCCAAGTGTCACACTCCAA 779

Qy     830 GGGCCACCAGAAGGCAGGGCCCTCAAGACGACCAGTCAGTCTTTTCTTTGC 889
           ||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     780 GGGCCACCAGAAGACGCGCAAGGCCCTCAAGACCCACCACAGTCATCTGGCTTTTCTTCGC 839

Qy     890 CTGCTGGCTGCCATATTATGTGGGGATCAGTAACATCCTTTGGGAGTCAT 949
           ||||||||||||||||||||||||||||||||| |||||||||||| ||||
Db     840 CTGTTGGCTGCCTTTACTACACTATTGGGATGGATCACTGACATCGACTCGACTCGACTGTCATCTTCCTCGAAATCAT 899

Qy     950 CAAGCAAGGATGTGACTTCGAGAGCATGTGCACAAGTGGATCTTCCATCACAGAGGGCCT 1009
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||| |||||
Db     900 CAAGCAAGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGAATTCCATCACCGAGCCCT 959

Qy    1010 CGCCTCTTCCACTGTTGCCTGAACCCATCCTCTATGCTTCCTCGGGCCAAGTTCAA 1069
           |||||||||||||||||||| ||||||||||||| |||| |||||||| |||||||
Db     960 AGCTTCTTCCACTGTTGTCTGAACCCCATCCTCTATGCTTTCCTCTGGGAGCGCAAATTTAA 1019

Qy    1070 AAGCTCTGCCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCCAGATCCTTTTC 1129
```

RESULT 3

```
AAQ80521
ID  AAQ80521 standard; cDNA; 1737 BP.
XX
AC  AAQ80521;
XX
DT  25-MAR-2003  (revised)
DT  18-JUL-1995  (first entry)
XX
DE  Human monocyte PF4AR cDNA.
XX
KW  Interleukin-8 receptor; IL-8 receptor; PF4AR;
KW  platelet factor superfamily receptor; monocyte; chemotactic;
KW  inflammation; inflammatory disease; arthritis; emphysema; cystic;
KW  fibrosis; colitis; bronchitis; meningitis; therapeutic; ss.
XX
OS  Homo sapiens.
XX
FH  Key             Location/Qualifiers
FT  CDS             91...1149
FT                  /*tag= a
XX
PN  WO9428931-A1.
XX
PD  22-DEC-1994.
XX
PF  07-JUN-1994;  94WO-US006380.
XX
PR  11-JUN-1993;  93US-00076093.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Chuntharapai A, Lee J, Hebert C, Kim KJ;
XX
DR  WPI; 1995-036114/05.
DR  P-PSDB; AAR68812.
XX
PT  Treatment of inflammatory disorders - by administering an antibody
PT  capable of binding a platelet factor 4 superfamily receptor polypeptide.
XX
PS  Disclosure; Page 54-56; 83pp; English.
XX
CC  2 PF4AR members were identified by probing lambda libraries from human
CC  monocyte-like cell line HL-60 and human peripheral blood lymphocytes
CC  using a large fragment of IL-8 receptor DNA (full sequence given in
CC  AAQ80520). The nucleotide sequences of the 2 PF4ARs are given in AAQ80521
CC  and AAQ80522, and their respective amino acid sequences in AAR68812 and
CC  AAR68813. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ  Sequence 1737 BP; 454 A; 411 C; 373 G; 499 T; 0 U; 0 Other;
```

```
Qy    1620 TGCTGTGTGATGGTTTGTTTGGTTTTTTTTTTTGTTGTTGTTGTTTTTTTTTTTTCTGTAA 1679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 TGCTGTGTGATGGTTTGTTTGGTTTTTTTTGTTTGTTGTTGTTTTTTTTTTTTCTGTAA 1560

Qy    1680 AAGATGGCACTTAAAACCAAAGCCTGAAATGGTGGTAGAAATGCTGGGGTTTTTTTTGTT 1739
           ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    1561 AAGATGGCACTTAAAACCAAAGCCTGAAATGGTGCTAGAAATGCTGGGGTTTTTTTTGTT 1620

Qy    1740 TGTTTGTTTTTTCAGTTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTT 1799
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 TGTTTGTTTTTTCAGTTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTT 1680

Qy    1800 GTATTACAATGTGTAATAAAAGTCAATGATAAAACTTAAAAAAAAAAAAAAAAAAAAAAA 1859
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GTATTACATTGTTAATAAAAGTCAATGATAAAACTTAAAAAAAAAAAAAAAAAAAAAAA 1740

Qy    1860 AAAAAAAAAAAAAAAA 1877
           ||||||||||||||||
Db    1741 AAAAAAAAAAAAAAAA 1758
```

```
XX
PN   US5440021-A.
XX
PD   08-AUG-1995.
XX
PF   25-FEB-1994;   94US-00202056.
XX
PR   29-MAR-1991;   91US-00677211.
XX
PA   (HEBE/)  HEBERT C.
PA   (CHUN/)  CHUNTHARAPAI A.
PA   (KIMK/)  KIM K J.
PA   (LEEJ/)  LEE J.
XX
PI   Chuntharapai A,   Kim KJ,   Hebert C,   Lee J;
XX
DR   WPI; 1995-283151/37.
DR   P-PSDB; AAR80757.
XX
PT   New antibodies against interleukin 8 type B receptor - used to treat or
PT   prevent inflammation, also for detecting receptor expression and
PT   purificn.
XX
PS   Example 2; Col 47-50; 62pp; English.
XX
CC   Antibodies directed against the interleukin-8 receptor B can be used to
CC   treat or prevent inflammation e.g. psoriasis, dermatitis, rheumatoid
CC   arthritis and particularly inflammatory bowel disease and chronic lung
CC   inflammation. When immobilised, these antibodies may be used to detect
CC   interleukin-8 receptor B expression in cells and tissues and for affinity
CC   purification of interleukin-8 receptor B from cells. This sequence is an
CC   additonal chemokine superfamily receptor which was identified by probing
CC   lamda libraries of genomic DNA from a human monocyte-like cell line (L-
CC   60) and human peripheral blood lymphocytes using a large fragment of the
CC   interleukin-8 type A receptor DNA (See AAQ99006). (Updated on 25-MAR-2003
CC   to correct PF field.)
XX
SQ   Sequence 1737 BP; 454 A; 411 C; 373 G; 499 T; 0 U; 0 Other;

Query Match            51.9%;  Score 974.4;  DB 2;  Length 1737;
Best Local Similarity  76.1%;  Pred. No. 2.9e-178;
Matches 1391;    Conservative  0;   Mismatches 301;   Indels 136;   Gaps 10;

Qy    50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC 109
         |||||  || |||||||  ||||||  ||||| ||||| |||| |||||| |||| ||||
Db    20 GCGCGGCGGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTGCTCCACCGCCATCTGAGAGAAC 79

Qy   110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAATACTACTGAAGA 169
         |||  |||||||||||      |||||||||| |||||||||| ||||||||||| || ||
Db    80 CAGCGGTTACCATGGA-----GGGGATCAGTATATATACACTTCAGATAATACTACCGAGGA 134

Qy   170 AGTGGGGTCTGGAGACTATGATCAACAAGGAGACCCTGCTTCCGGGATGAAAACGTCCA 229
         | ||||| || ||| ||||||||  |  || || |||| || || || ||||| | ||
Db   135 AATGGGGCTCAGGGGACTATGATCCATGAGGGGAACCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy   230 TTTCAATAGAGATCTTCCTGCCCACCATCACTTCCTCCATCATCTTCCTCCATCATCTTCTTAACTGGCATTGTGGG 289
         |||||||  | ||| |||||||||||||||  |||| ||||||  |||||  ||||| |||||||| |||||||||||
Db   195 TTTCAATAAAATCTTCCTGCCCACCATCACTTCCTCCATCATCTTCCTCCATCATCTTCTTAACTGGCATTGTGGG 254

Qy   290 CAATGGATTGGTGATCCTGGTACCAGAAGAGCTAAGGAGGCATGACGGACAA 349
         |||||||||| ||||||||||||||||||||||||||||||| ||| |||||
Db   255 CAATGGATTGGTGATCCTGGTACCAGAAGAGCTGAAGAACTGAAGACGGACAA 314

Qy   350 GTACCGGCTGACCCGTGTCAGTGGGCTGACCGTCTGATAATACACTCCCCTTCTGGGC 409
         |||||||||||||||   ||||| |||  || |||||||  | || |||||||||||||
Db   315 GTACACAGGCTGACCCGTGTCAGTGGGGTTGATTTCAGCACCTACACGCTTCCCTTTCTGGGC 374

Qy   410 AGTTGATGCCATGGCTGACTGGTACTTTGTGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
         ||  |||||   |||  | |||||| |||||||| ||||||||||| || |||||||||||
Db   375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy   470 CTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTA 529
         |||| ||||||| ||||||||||| |  |||||||| ||  || |||||||||||| |||
Db   ...
```

```
Db  1020 AACCTCTGCCCAGACACGCACTCACCTCTGTGAGCAGAGGGTCCAGCCTCAAGATCCTCTC 1079
          |||||||||||| ||  |||||||| |||||||||||||||||||||| ||||||| |||
Qy  1130 CAAAGGAAAGCGGGGTGGACACTCCTTCCGTCCCACGGAGTCAGAATCCTCCAGTTTCA 1189
          |||||||||||| | ||||| ||  ||||  ||| |  ||| ||| ||||||||||||||
Db  1080 CAAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGCTGAGTCTTCAAGTTTTCA 1139
Qy  1190 CTCCAGCTAACCCTTATGCAAGACTTATATATATATATATATATATAAGAACTT 1249
          ||||||||||       ||| |||  |||||||||||  |||||||| |||| ||||
Db  1140 CTCCAGCTAA-------CACAGATGTAAAAGACTTTTTTTTTTATACGATAAATAACTTTT 1191
Qy  1250 TTTTATGTTACACATTTCCAGATAATAAAGAGACTGACCAGTCTTCGTACAGTTTTTTTT 1309
          ||||  |||||||||| |||||||||||| | ||||||||  |||||||||||||||  ||
Db  1192 TTTTAAGTTACACATTTTCAGATAATAAAAGACTGACCAATATTGTACAGTTTTTATTGC 1251
Qy  1310 TTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTTGACTTAATTT 1369
          ||  |   ||||  ||| ||  || ||||||| |||||||||||||| |||  |||||| |
Db  1252 TTGTTGGATTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATTGACTTA 1305
Qy  1370 ATATAAATATTGTTTTTTGTTGTTTCATGTGAATGAGCGTCTAGGCAGCAGGACCTGTGGCC 1429
          |||| |  |||| ||||| |||||||||| ||| || | |||||||||| |||||||||| ||
Db  1306 TTTATA---TAAATTTTTTTGTTTCATATTGATGTGTGTCTAGGCAGGACCTGTGGGCC 1361
Qy  1430 AAGTTCTTAGTAGCTGTTTATCTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGA 1489
          |||||||||| ||||||   | |  || |||||||||||       ||||| ||||||
Db  1362 AAGTTCTTAGTTGCTGTATGCTCCGTGGGTAGGACTGTAGAA-------AAGGGAACTGA 1413
Qy  1490 ACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGTGTTGCTGCA 1549
          |||||||||| | ||  |   |  || ||||||||||  ||||||  |||||| || |||
Db  1414 ACATTCCAGAGCGGTGAGTGAATCACGTAGAAGCTAGAAATGATCCCCAGCTGTTTATGCA 1473
Qy  1550 TAATCTTCATTCCGGAGGAGGAGGAGCACCCCCACCCCCACCCCCATTCTTAAAT 1609
          || |    |||  ||
Db  1474 TAGATAATCTCCTCCA----------------------------------- 1488
Qy  1610 TGTTTGGTTATGCTGTGTGATGGTTGTTGTTGGTTTTTTTTTGTTGTTGTTTTTT 1669
                              |||||  |||  |||||| |||  |||| ||||||
Db  1489 -------------------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGAT 1525
Qy  1670 TTTTCTGTAAAAGATGGCACTTAAAAACCAAAGCCTGAAAATGGTGGTAGAAATGCTGGGGGT 1729
          ||| |||||| |||||||||||| ||||||||||  || | | ||||  ||||||||||||
Db  1526 TTTGCTGTAGAAGATGGCACTTATAACCAAAGCCCAAAGTTGGT-ATAGAAATGCTGG--- 1581
Qy  1730 TTTTTTTGTTTTGTTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCTACAAAT 1789
                        ||||||||||||||| ||||| |  |||| |||||| ||||  |||
Db  1582 --------------TTTTTCAGTTTTCAGGAGTGGGTTGATTTCAGCACCTAC-AGT 1623
Qy  1790 GTACAGTCTTGTATTACATTGTTAATAAAGTCAATGAATTAAAAAAAAAAAAAAA 1849
          |||||||||||||||  |||||||||||||| ||||  ||||| |||||||||||||||
Db  1624 GTACAGTCTTGTATTAAGTTGTTAATAAAGTACATGTTAAACTTAAAAAAAAAAAAAAAA 1683
Qy  1850 AAAAAAAAAAAAAAAAAAAAAAAAA 1877
          ||||||||||||||||
Db  1684 AAAAAAAAAAAAAAAAAAAAAAA 1711
```

```
RESULT 4
AAQ99007
ID   AAQ99007 standard; cDNA; 1737 BP.
XX
AC   AAQ99007;
XX
DT   25-MAR-2003  (revised)
DT   26-MAR-1996  (first entry)
XX
DE   Chemokine superfamily receptor coding sequence.
XX
KW   Interleukin; IL-8; inflammation; psoriasis; dermatitis;
KW   rheumatoid arthritis; inflammatory bowel disease;
KW   chronic lung inflammation; treatment; antibody; affinity purification;
KW   detection; ss.
XX
OS   Homo sapiens.
```

```
Db   435  CTACACAGTCAACCTCTACAGACGTGTCCTCATCCTGCTGCCTTCATCAGTCTGCGACCGCTA  494
Qy   530  CCTCGCCATTGTCCAGGCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGC     589
                ||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db   495  CCTCGGCCATCGTCCACGGCCACCAACAGTCAGAGGCCACCAACAGTCAGGAGAGCTGTTGGCTGAAAAGGT  554
Qy   590  AGTCTATGTGGGCGTCTGGATCCACGCCCTCCTGCTGACTATACCTCATCTTTGC  649
               |||||||||| ||||||||||||||||||||||||||||||||||
Db   555  GGTCTATGTTGGCGTCTGGATCCGCTCCTCCTGTGCTGACTATTCCCGACTTCATCATTTGC  614
Qy   650  CGACGTCAGCCAGGGGGACCATCAGTCAGGGGATGACAGGTACATCTGTGACCGCCCTTTA  709
               |||                 |||||||||||||||||||||||||||||||||||||
Db   615  CAACG------------TCAGTGAGGCAGCAGATGACAGATATCTGTGACCGCTCTA  659
Qy   710  CCCCGATAGCCTGTGGATGGTGGTTGTTTTCAATTCCAGCATTCCAGCATTCCAGTTTCATCCT  769
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660  CCCCAATGACTTGTGGGTGGTTGTTCCCAGTTTCAGCATTTCAGTTTGGCCTTATCCT  719
Qy   770  GCCCGGCATCGTCATCCTCCGTTACTGACTCATCATCATCTCTAAGCTGTCACACTCCAA  829
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720  GCCTGGTATTGTCATCCTCCGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAA  779
Qy   830  GGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTTCTTTGC  889
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780  GGGCCACCAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTCGGCTTTCTTCTCGC  839
Qy   890  CTGCTGGCTGCCCATATTATGTGGGGATCCAGCATCGACTGCCTCCTTCATCCTTTGGGAGTCAT  949
               |||||||||||       ||||||||||||||||||||||||||||||||||||||
Db   840  CTGTTGGCTGCCTTACTACATTGGGATCAGCATCAGCATCGACTTCCTTCATCCTTCATCTTGGAAATCAT  899
Qy   950  CAAGCAAGGATGTGACTTCGAGAGCATTGCACAAGTGACTCCATCCACAGAGGGCCCT  1009
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db   900  CAAGCAAGGTGTGAGTTTGAGGACACTGTGCACAAGTGATTTCCATCATCACCGAGGGCCCT  959
Qy  1010  CGCCCTTCTTCCACTGTTGCCTGAACCCCATCCTTCCTTCCTCGGGGCCAAGTTCAA  1069
               ||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   960  AGCTTTCTTCCACTGTGTCTGAACCCCATCCTTCCTTCCTTGGAGCCAAATTTAA  1019
Qy  1070  AAGCTTCTGCCCAGCAGTGCACTCAACTCACCTCCATGAGCAGAGGCTCCAGCGCCTCAAGATCCTTTC  1129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020  AACCTTCTGCCCAGCAGTGCACGCACGCTCACCTCCTGTGAGCAGAGAGGTCCAGCGCCTCAAGATCCTCTC  1079
Qy  1130  CAAAGGAAAGCGGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTTCA  1189
               |||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  1080  CAAAGGAAAGCGAGGTGGACATTCCACTGTTCCCACTGAGTCTGAGTCTTCAAGTTTCA  1139
Qy  1190  CTCCAGCTAACCCTTATGCAAAGACTTATTATAATAATAATATATATAGAAAGAACTT  1249
               ||||||||||   ||| |                 |||||||||||||||
Db  1140  CTCCAGCTAA------CACAGATGTAAAAGACTTTTTTTTATACGATAAATAACTTTT  1191
Qy  1250  TTTTATGTTACACATTTTCAGATATAAGAGAGACTGACCAGTCTTGTGTACAGTTTTTTTTT  1309
               |||||||||||||||||||||||||||| |||||||||| ||||||||||||||
Db  1192  TTTTAAGTTACACATTTTCAGATATAAAGACTGACCAATATTGTACAGTTTTTTTTATTGC  1251
Qy  1310  TTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTTGACTTAATT  1369
               ||||||||||||| |                |||||||||||||||||||||
Db  1252  TTGTTGGATTTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATTGACTTA  1305
Qy  1370  ATAAATAATATTGTTTTTTTGTTTTGTTTTCATGTGAATGAAGTGAGCCTGTAGGCAGACCTGTGGCC  1429
               ||||| |       ||||||||||||||||||||||||||||||||||||||||||
Db  1306  TTTATA---TAAATTTTTTTTTGTTTCATATTGATGTGTGCGTAGGCAGGACCTGTGGGGC  1361
Qy  1430  AAGTTCTTAGTAGCTGTTGTTTATCTGTGTGTAGGACTGTGAACTGTAGAACTGTGAGAGAGAAACTGA  1489
               ||||||||||||||||||||||||||||||||||||||||||||        |||||||||
Db  1362  AAGTTCTTAGTTGCTGTTGCTGTATGTCTCGTTGGTGGGACTGTAGAA-------AAGGGAACTGA  1413
Qy  1490  ACATTCTTCAGAATGTGTGGTAAATTGAATAAAGTACGCCGTGATCCTCAGCTGTTCTGCTGCA  1549
               ||||||||||||||||   ||||||    |||||||||||||||||||||||||||||
Db  1414  ACATTCTTCCAGAGCGTGTAGTGAATCACGTAAAGCTAAGAAATGATCCCCAGTGTGTTTATGCA  1473
Qy  1550  TAATCTTCTTCATTCCGAGGAGCACCCCCACCCCCCCACCCCCCCATTCTTAAAT  1609
               |||           ||||||||||||||||||||||||||||||||||||||
Db  1474  TAGATAAATCTCTCCA---------------------------------------  1488
```

```
Qy  1610  TGTTTGGTTATGCTGTGTGATGGTTTGTTGTTTTTTTGTTTGTTGTTGTTTT  1669
               |||||||||||||||||||||                    |||||
Db  1489  -----------------------TTCCCGTGGAACGTTTTTTCCTGTGTTCTTAAGACGTGAT  1525
Qy  1670  TTTTCTGTAAAAGATGGCACTCAAAACCAAAGCCTAAAACCAAAGGTGGTGGTAGAAATGCTGGGGT  1729
               ||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db  1526  TTTGCTGTAGAAGATGGCACTTATAACCAAAGCCCAAAGTGGT-ATAGAAAATGCTGG---  1581
Qy  1730  TTTTTTTGTTTGTTGTTTTTTTCAGTTTCAGTTTTTCAAGTAGTTGACTTCAGTCCTACAAAT  1789
               |||||||||||||                    ||||||||||||||||||
Db  1582  -----------------TTTTTCAGTTTCAGGAGTGGGTTGATTTCAGCACCTAC-AGT  1623
Qy  1790  GTACAGTCTTGTGATTACATTGTTAATAAAAGTCAATGATAGATAAACTTAAAAAAAAAAAAA  1849
               |||||||||||||||||||||||||||||||||||||||||||
Db  1624  GTACAGTCTTGTGATTAAGTTGTTAATAAAGTACATGTTAAACTTAAAAAAAAAAAAAAA  1683
Qy  1850  AAAAAAAAAAAAAAAAAAAAA  1877
               |||||||||||||||
Db  1684  AAAAAAAAAAAAAAAAAAAAAA  1711
```

RESULT 5
```
AAQ29506
ID   AAQ29506 standard; DNA; 1737 BP.
XX
AC   AAQ29506;
XX
DT   25-MAR-2003  (revised)
DT   12-MAR-1993  (first entry)
XX
DE   New platelet factor 4 receptor superfamily member PF4ARi.
XX
KW   IL-8R; G-protein coupled receptor family; rhodopsin superfamily;
KW   pro-inflammatory cytokine; 8rr.20.15; ss.
XX
OS   Homo sapiens.
XX
FH                Location/Qualifiers
FT   CDS          91..1149
FT                /*tag= a
FT                /product= "PF4ARi"
XX
PN   WO9217497-A1.
XX
PD   15-OCT-1992.
XX
PF   23-MAR-1992;   92WO-US002317.
XX
PR   29-MAR-1991;   91US-00677211.
PR   19-DEC-1991;   91US-00810782.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Lee J, Holmes WE, Wood WI;
XX
DR   WPI; 1992-366191/44.
DR   P-PSDB; AAR27792.
XX
PT   Isolated human platelet factor 4 super-family receptor polypeptide and
PT   corresp. antibodies and DNA - useful as diagnostic and screening agents,
PT   and for treating inflammation or PF4AR-mediated disorders.
XX
PS   Claim 7; Fig 4; 78pp; English.
XX
CC   The IL-8 receptor cDNA sequence was isolated (see AAQ29505) and a 874bp
CC   sub-fragment of the coding sequence was used as a probe to screen human
CC   cell line HL60 and human peripheral blood lymphocyte cDNA libraries. Two
CC   new gene sequences were found that are clearly related to the IL-8
CC   receptor. One of these was contained in combined clone 8rr.20.15 and is
CC   predicted to encode an amino acid sequence which is 34% identical with
CC   both the high and low affinity IL-8 receptors. See also AAQ37107.
CC   (Updated on 25-MAR-2003 to correct PN field.)
```

XX
SQ   Sequence 1737 BP; 457 A; 412 C; 370 G; 498 T; 0 U; 0 Other;

Query Match   51.7%;   Score 969.6; DB 2;   Length 1737;
Best Local Similarity   75.9%;   Pred. No. 2.4e-177;
Matches 1388;   Conservative   0;   Mismatches 304;   Indels 136;   Gaps 10;

```
Qy         50 GCAGGTAGCAGTGACCCTCGAGGCGGTTGTGTCCGGTAACCACGGCTGTAGAGC 109
              ||||||||||| ||||| |||||||| |||||| |||||||||| || |||||||
Db         20 GCGCGGCGCAAAGTGACGCGGAGGCCTGAGCTGCTCCAGTAGCCACCGGATCCTGAGAAC 79

Qy        110 GAGTTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
              ||||| || ||||||||                  |||||||| |||||| |||||||||
Db         80 CAGCCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACACCGAGGA 134

Qy        170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCTGCTCCCGGATGAAAACGTCCA 229
              ||||||||||||||||||| |||||| ||||||||| ||||  |||||||| ||||||
Db        135 AATGGGCTCAGGGACTATGACTCCATGAGGAACCCTGTTCCGGAAGAAAATGCTAA 194

Qy        230 TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTGACTGGCATAGTCGG 289
              ||||||||| |||||||||||||||||||||||||||||||||| ||| |||| |||
Db        195 TTTCAATAAATCTTCCTGCCCACCATCTACTTGTCATCATCTTAACTGGCATTGTGGG 254

Qy        290 CAATGGATTGGTGATCCTCGTCATGGGTTACCAGAAGAAGCTAAGGACGTAACGGACAA 349
              ||||| |||||| |||||||||||||||||||||||||||| |||||||||||||||||
Db        255 CAATGGATTGGTCATCCTCGGTCATGGGTTACCAGAAGAAGCTGAGAAGCATGACGGACAA 314

Qy        350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTTCCCCTTCTGGGC 409
              |||| |||| |||||||||||||||||||||||| |||| || |||||||||||||||||
Db        315 GTACAGCTGCTGCACCTGTCAGTGGCGCGACCTCCTCTTTGTCATCACGCTTCCCTTCTGGGC 374

Qy        410 AGTTGATGCCATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
              ||||||||||| ||||| |||||||||||||| ||| ||||| |||||||||||||||
Db        375 AGTTGATGCCGTGGCAAACTCGTACTTTGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy        470 CTACACTGTCAACCTCTACAGCAGCGTTCATCAGCCTGGCACTTCATCAGCCTGGACCGGTA 529
              ||||||||||||||||| ||||||||| ||||||||||||||||| |||||||||||||
Db        435 CTACACAGTCAACCTCTACAGCAGCGTCCCTCACGCCTCGCCTTCACCCGACCGCTA 494

Qy        530 CCTCGCCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGAACTGCTGGCTGAAAAGGC 589
              |||| ||||| || |||||||||||||| |||||||||||| || |||||||||||||
Db        495 CCTGGCCATCGTCCAGCCACCAACAGTCAGAGGCTGTTGGCTGAAAAGGT 554

Qy        590 AGTCTATGTGGGCGTTCTGGAGATCCCAGGCCCTCCTCGACTATACCTGACTTCATCTTTGC 649
              |||| ||||| |||||||| ||| |||||||||||| |||||||||||||| |||||||||
Db        555 GGTCTATGTTGGCGTCTGGATCCCTGCCTCCTCGCTGCTGACTTCATCTTTGC 614

Qy        650 CGACGTCAGCAGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCCTTTA 709
              ||||||||||||||||||| || ||||||||| || ||||||||||||||||||||||||
Db        615 CAACG---------TCAGTGAGGCAGATGACAGATATCTGTGACCGCCTTCTA 659

Qy        710 CCCCGATAGCCTGTGGATGGTCGTGTTGTTGTGTTCCAGTTTCCAGTTCCAGTTTCAGTGTT 769
              |||||||||||||| |||| ||||| |||||||||||||||||||||||||||||
Db        660 CCCCAATGACTTGTGGGGTGGTGGTGTTGTGTTCCAGTTTTCAGCACACATCATGTGGCCTTATCCT 719

Qy        770 GCCCGGCCATCGTCATCCTCCTGTTACTGACTCCATCATCTCCTAAGGCTGTCACACTTCCAA 829
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 GCCTGGTATTGTCATCCTGCTATTGCATTCATCTCCAAGCTGTCACACTCCAA 779

Qy        890 CTGCTGGCTGCCATATATGTGGGGATCAGCATCGACTCCTTCATCCTTTTGGGAGTCAT 949
              ||||| || ||| ||||| ||||||||||||||||||||||||||||  ||||||||||
Db        840 CTGTTGGCTGCCTTACTACATTGGGATCCATTGGACCTCCATCCTCCTGGAAATCAT 899

Qy        950 CAAGCAAGGATGTGACTTCGAGAGCATTGTGCACAAGTGGGATCTCCATCCATCAGAGGCGCCCT 1009
              ||||||||| ||| | |||||||||| |||||||| | ||| |||||||||| |||| || |||
Db        900 CAAGCAAGGGTATGAGTTTGAGAACACTGTGCACACTGTGGATTTCCATCACCACCGGAGGCCCT 959

Qy       1010 CGCCTTCTTCCACTGTTGCCTGACCGAACCCCATCCTCTATGCCTTCCTCGGGGCCAAGTTCAA 1069
              |||||||||||||||||||| |||| |||||||||||||||||||||||||||||| || |
```

```
Db        960 AGCTTTCTTCCACTGTTGCTGGTGAACCCCATCCTCTATGCATTCCTTCTCTATGCATTCCTTGAGCCCATCAAGATCCTTAA 1019
              ||||| || |||||||||||| |||| |||||||||||||||||||||||||||| ||||||||
Qy       1070 AAGCTCTGCCCAGCATGCATCCATGAGCAGCAACTCCATGAGACAGCGGCTCCAAGATCCTTTC 1129
              |||||||||||||||||||| |||||||||| ||||| |||||||||||||||||||| ||||
Db       1020 AACCTCTGCCCAGCACGCACTCACCTCTGTGACGCAGAGGTCCAGCCTCAAGATCCTCTC 1079
              ||||||||||||| || ||| ||||||||||| ||| |||||||||||| |||||||||
Qy       1130 CAAAGGAAAGCGGGGTGGACACTCTTCCGCTCTCCACGGAGTCGACAATCCTCCAGTTTTCA 1189
              |||||||||||||||||||||||||||||||||||| |||||||| ||| ||||||||||||
Db       1080 CAAAGGAAAGCAGGTGGACATTCATCTGTTCCAC-GAGTCTGAGTCTTCAAGTTTCA 1139
              ||| || |||||||| |||||| ||||  ||||||||||||||| ||| ||||||||||
Qy       1190 CTCCAGCTAACCCTTATGCAAAGACTTATATAATATATATATATATGATAAAGAACTT 1249
              |||||||||     |||||||||||||||||||| ||||||  ||||||||| |||||
Db       1140 CTCCAGCTAA-------CACAGATGTAAAAAGAACTTTTTTTTATACGATAAATAACTTTT 1191
              |||||||||          ||| |||| |||||||||||||||||||| ||||||||||
Qy       1250 TTTTATGTTACACATTTTCAGATATAGAGACTGACCAGTCTTGTACAGTTTTTTTTT 1309
              ||||| ||||| |||||||||||||||| |||||||||||||||||||| |||||||
Db       1192 TTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTTTATTGC 1251
              |||| |||||||||||||||||||||   ||||||||||| || |||||||||||
Qy       1310 TTTTAATGACTGTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGACTTAATTT 1369
              |||||||||||| ||||  ||||||| || |||||| ||| |||||| |||| |||
Db       1252 TTGTTGGATTTTGT-----CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAAATTGACTTA 1305
              ||| ||||||||       ||||||| ||||||||||| ||| |||||||||||||||
Qy       1370 ATATAAATATTGTTTTTGTTGTTTTCATGTGAATGAGCGTCTAGGCAGGACCGTGTGGCC 1429
              ||||| |||| ||| ||| ||||||||||||   |||| ||||||||||||||||||||||
Db       1306 TTTATA---TAAATTTTTTGTTTCATATTGTGTCGTAGGCAGGACCGTGTGGCC 1361
              |||||       |||||| |||| |||| |||||||||||||||||| |||||||
Qy       1430 AAGTTCTTAGTAGCTGTTTATCTGTGTGTAGGACTGGTAGGAACTGTAGAGGAAGAAACTGA 1489
              |||||||||||| |||||||||| ||||||||||||||| || |||||      ||||||||
Db       1362 AAGTTCTTAGTTGCTGTGTATGTCTTCGTGGTAGGACTGGTAGGA--------AAGGAACTGA 1413
              |||||||||||||||||| || || ||||||||||||||||           ||||||||
Qy       1490 ACATTCCAGAATGTGTGGTAAATTGAATAAATGAAAGCTAGCCGTGATCCTCAGCTGTTCTGCA 1549
              |||||||||||||||||||||| |||||| ||| |||||||||||||||| ||||||| |||
Db       1414 ACATTCCAGAGCGTGTAGTGAATCACGTAAAGCTAGAAATGATCCCCAGTCGTGTTTATGCA 1473
              ||||||||||| |||| |||||| || ||||||| ||||||||||||||  ||||||||||
Qy       1550 TAATCTTTCATTCCGAGGAGCACCCCACCCCACCCCACCCCATCCCATTCTTAAAT 1609
              |||| ||| ||||||||
Db       1474 TAGATAATCTCCA---------------------------------------------- 1488
Qy       1610 TGTTTGGTATGCTGGTGTGTGATGGTTTGTTTTGGTTTTTTTTTTGTTGTTGTTTTTTTT 1669
              ||||||||||||||||||||||||||||||| ||  ||||||||||   ||||||||||
Db       1489 -----------------------------------TTCCCGTGGGAACGTTTTTCCTGTTCTTAAGACGTGAT 1525
              |                                   ||||||||||| |||||||||||||
Qy       1670 TTTTCTGTAAAAGATGGCACTTAAAAACCAAAGCCTGAAAATGGTGGTAGAAATGCTGGGGT 1729
              |||| ||||||||||||||| |||||||||||||||||||||||| ||| |||||| |||||
Db       1526 TTTGCTTGGAAGATGGCGACTTATACCAAGGGAT-ATAGAAATGTGTGC--- 1581
              ||| | ||| ||||||||| ||||||||||||    |||||| |||
Qy       1730 TTTTTTTGTTTGGTTGTTTTTTTCAGTTTTCAGTTTGACTTCAGTCCCTACAAAT 1789
              |||||||||||||||||||||| ||||||    |||||||||||| |||||||
Db       1582 ----------TTTTTCAGTTTTCAGGAGGTGGTTGATTTTCAGCACCTAC-AGT 1623
              ||||||||||||||||| ||| || ||||||| ||| ||||
Qy       1790 GTACAGTCTTGTATTACATTGTAATAAAAGTCAATGAAAGTTAAAAAAAAAAAA 1849
              ||||||||||||||||||| |||||||||| |||  ||||||||||||||||||
Db       1624 GTACAGTCTTGTGTATTAAGTTGTTAATAAAGTACATGTTAAACTTAAAACTTAAAAAAAAAAAAAAA 1683
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||
Qy       1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
              |||||||||||||||||||||||||||||
Db       1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

RESULT 6
AAD12801
ID   AAD12801 standard; DNA; 1679 BP.
XX
AC   AAD12801;
XX
DT   16-OCT-2001 (first entry)
XX
DE   Human neuropeptide Y (NPY) Y3 receptor DNA.
XX
KW   Human; neuropeptide Y; NPY; bone disease; bone mass; gene therapy;
KW   cerebrospinal fluid; CSF; inositol phosphate; IP; Paget's disease;
KW   fracture; extracellular signal-regulated kinase; ERK; osteoporosis;
KW   osteopenia; bone metastasis; neurotransmitter; osteogenic;

```
KW   NPY Y3 receptor; ds.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            89..1147
FT                  /*tag= a
FT                  /product= "Human neuropeptide Y (NPY) Y3 receptor"
XX
PN   WO20015477-A1.
XX
PD   26-JUL-2001.
XX
PF   22-JAN-2001; 2001WO-US002040.
XX
PR   20-JAN-2000; 2000US-00489872.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
PA   (AMLI/) AMLING M.
XX
PI   Amling M, Karsenty G, Ducy P;
XX
DR   WPI; 2001-488709/53.
DR   P-PSDB; AAE06690.
XX
PT   Treating or preventing reduced bone mass, e.g. osteoporosis, by reducing
PT   the level of neuropeptide Y activity in blood or cerebrospinal fluid.
XX
PS   Example 7; Page 84-85; 102pp; English.
XX
CC   The present invention relates to a method for treatment or prevention of
CC   bone diseases characterised by loss of bone mass, comprises administering
CC   to a mammal a compound that lowers the level of neuropeptide Y (NPY) in
CC   the serum or cerebrospinal fluid (CSF) or a compound that lowers the
CC   level of inositol phosphate (IP) or extracellular signal-regulated kinase
CC   (ERK). The method is specifically used to treat (including by gene
CC   therapy) or prevent osteoporosis, osteopenia or Paget's disease. These
CC   diseases may also be diagnosed by detecting elevated NPY levels,
CC   also be used e.g., in cases of fractures or bone metastases. These
CC   including monitoring of treatment, assessing efficacy of compounds in
CC   clinical trials and for identifying subjects at risk. The present
CC   sequence is a human NPY Y3 receptor DNA
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match             50.8%;  Score 952.8;  DB 4;  Length 1679;
Best Local Similarity   80.2%;  Pred. No. 4.1e-174;
Matches 1210;  Conservative  0;  Mismatches 252;  Indels  46;  Gaps  6;
```

```
Qy    44  GCAGGTGCAGGTAGCAGTGACCCCTCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTG   103
              ||||||||||||| |||  ||||  |||||||||| || ||||||| ||||| ||||
Db    12  GCGGGCAGCAGGTAGCAGAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG   71

Qy   104  TAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATCACTTCTGATAACTACTC   163
              ||||| |||||||||||| || |||||| ||||| ||||||||||||| || |||||||||||
Db    72  GAGAACCAGGGGTTACCATGGA----GGGGATCAGTATATACACTTCAGATAACTACAC   126

Qy   164  TGAAGAAGTGGGGTCTGGACTGTGGACTACGATGATGAACCCGCTCCGGGATGAAAA   223
              ||||| ||||||| |||||| ||||||| ||| |||||||||||||||||||||| |||||
Db   127  CGAGGAAATGGGCTCAGGGACCATGAGCACTTCCAGTGATGAGAATTCCGTGAAGAAAA   186

Qy   224  CGTCCATTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCAT   283
              ||||| || |||| || |||||| |||||||||||||| |||| ||||||| ||||||
Db   187  TGCTAATTTCAATAAAATCTTCCTGCCCACCATCCACCATCTTCTTAACTCGGCAT   246

Qy   284  AGTCGGCAATGGATTGGTGATCCGGTCATGGGTCATGTCGTCGTGCGTTACCAGGAAGTCAAGGAGCATGAC   343
              |||||| ||||| |||||| ||  ||||| ||||||| ||||||||| ||||||| |||||||||||||||
Db   247  TGTGGGGCAATGGATTGGTCATCCTGTCATGGGTCATGTCGTCATGGGTTACCAGAAAGCATGAC   306

Qy   344  GGACAAGTACCGGCCGCACCTGCTCAGTGTGACTGCCACCTTGTCATCACACTCCCTT   403
              ||  |||||| |||||| || |||||||||| |||| |||||||||||| ||| |||||
Db   307  GGACAAGTACAGGCTGCACCTGGTCAGTGGCCGACCCTCCCTTTGTCATCACGCGCTTCCTT   366
```

```
Qy   404  CTGGGGCAGTTGATGCCATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGTCCA   463
              ||||||||||||||||||||||||||||||||||||||||||   || ||| ||||
Db   367  CTGGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAGCAGTCCA   426

Qy   464  TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGA   523
              | ||||||||||||||| || |||| |||||||||| ||||||||||||||||||||||
Db   427  TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCCTGGCCTTCATCAGTCTCGGA   486

Qy   524  CCGGTACCTGCCATTGTCCACCAGTCAGTCGATCAGTCAAGGGCCAAGGAAACTGCTGCTGA   583
              || |||||||| ||||| || |||||| ||||||| ||||||| |||||| |||||||||
Db   487  CCGCTACCTGGCCATCGTCCACCGCCACCAGTCAGTCCAACAGTCAGAGCGAAGCGTGTTGGCTGA   546

Qy   584  AAAGGACAGTCTATGTGGGCCTCTGATCCGACCCTCCCTCCGACTATACCTGACTTCAT   643
              | ||||| ||||| ||||| ||| ||| ||| ||| || |||||| ||||||||||||||
Db   547  AAAGGTGGTCTATGTTGGGTGGTTGGGGTCGATCCCCGTCCTTGGTGTGACTTCCGGACTTCAT   606

Qy   644  CTTTGCCGACGTCAGCCAGCAGGGGACATCAGCAGGGGGATGACAGGTACATCTGTGACCG   703
              || ||||||||           ||||||||| |||||||||||||||||
Db   607  CTTTGCCCACG-----------TCAGTGAGGCAGATGACAGATATCTGTGACCG   651

Qy   704  CCTTTACCCGATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCT   763
              || ||||||| || |||||||||| ||||| |||||||||||||| ||||||||||||||
Db   652  CTTTACCCCCAATGACTTGTGGTGTGGGTGTTGTTCCAGTTTCAGCACATCATGGTTGGCCT   711

Qy   764  CATCCTGCCCGCATCGTCATCCTCTCCTGTTACTGCATCATCATCCTCTAAGCTGTCACA   823
              ||||||||  |||||||  ||||||||||||| |||||||||||||||| ||| |||||||
Db   712  TATCCTGCTCGTGATTGTCATTCTGCTATTCCATTATCATCCTCCAAGCTGTCACA   771

Qy   824  CTCCAAGGGCCACCAGAAGCCAAGGCCCTCAAGACGCACAGTCATCCTCATCCTAGCTTT   883
              |||||||||| || |||| ||||||||| ||||| ||| | ||| |||||||| |||||||
Db   772  CTCCAAGGGCCACCAGAAAGGCCAAGGCCCTCAAGACCACAGTCATCATCCTTCATCCTGGCTTT   831

Qy   884  CTTTGCCTGCTGGCTCCCATATTGGGGATCCAGATCAGCATGGGCATCGACTCCTTTTGGG   943
              ||||||||| ||||| |||||| ||||||||||| ||||| ||||||| |||||| |||
Db   832  CTTCGCCTGTGTGGCTGCCTTACTACATTGGGATCAGCATCCAGCATCCGATCATCTCCTGA   891

Qy   944  AGTCATCAAGCAAGGATGTGACTTCGAGAGATGTGCACAAGTGGACTCGATCTCCATCACAGA   1003
              |  ||||||||||||||||||||||  || || || ||||||||||||| || |||||||
Db   892  AATCATCAAGCAAGGTGTGAGTTTGAGAACACTGCGACACTGGGGATTTCCATCACCGA   951

Qy  1004  GGCCCTCCTTCTCCACTGTGCCTGACCCCATCCTCATGCGCTTCCTCATGCTTCCTCCAGCCAA   1063
              ||||| |||||| || |||||||| | |||| |||| ||||||||| |||||||||||
Db   952  GGCCCTAGCTTCTTCCACTGTTGCTGACCCCATCCTCATGCTCTGGAGCCAA   1011

Qy  1064  GTTCAAAAGCTCTGCCCAGCATGCACTCAACTCACTCATGAGAGAGGGCTCCAGCCTCAAGAT   1123
              || ||||||             |||||  ||||||||||| |||||||||| || |||||
Db  1012  ATTTAAAACCTCTGCCCAGCAGACGCACTCACCTCTGTCGTGAGCAGAGGGCCAGAATCCTCAAGAT   1071

Qy  1124  CCCTTTCAAAAGAAGAAAGCGGGGTGGACACTCTTCCGGTCTCCGCGAGTGCAGAATCCTCCAG   1183
              || |||||||||||||||| | ||||| |||||| |||||| |||| | ||||| |||| |||
Db  1072  CCTCTCCAAAGAAAGCGAGGTGGACATTCCATCTGTTTCCACTGTTTCACTGAGTGCTTCAAG   1131

Qy  1184  TTTTCACTCCAGCTAACCCTATGCAAAGACTTATATAATATATATATAA   1243
              ||||||||||||||||||||| |||||| ||||||||||      ||||||||||
Db  1132  TTTTCACTCCCAGCTAACAACAGATTTTTT------TTTATACGATAAATA   1183

Qy  1244  GAACTTTTTATGTTACACATTTCAGATAATAAGAAGACTGACCAGTCTGTACAGTTTT   1303
              ||||  |||| |||| ||||| |||||| || |||||| | ||| ||||| ||||||||
Db  1184  ACTTTTTTAAGTTACACATTTTCAGATAAAAAGACTGACCAATATTGTACAGTTTT   1243

Qy  1304  TTTTTTTTTTTAATTGACTGTTGGAGTTTTGTTTTTAGTTTTTGTTGAGTTTGACT   1363
              ||||||||||                    || |||  ||| |||| |||||||| |
Db  1244  TATTGCTTGTGGATTTTGT----CTTGTGTTGTTTGTTTTTCATGTGAAGTGTTTAATT   1297

Qy  1364  TAATTTATATAAAATATTGTTTTTGTTGTTTCATGTGAAGTGAGCGTCTAGGCAGGACCTT   1423
              ||||| || |||| |||||| |  |||||| ||||||| ||||||||||||||||||||
Db  1298  GACTTATTTATA----TAAATTTTTTGTTTTTCATATTGATGTGTGTCTAGGCAGGACCTT   1353

Qy  1424  GTGGCCAAGTTCTTAGTAGCTTGTTTATCGTGTGTAGGACTGTAGAGGAGGA   1483
              |||||||||||||||||||||||||||| ||||| || ||||||||||||| ||||
Db  1354  GTGGCCAAGTTCTTTAGTTGCTATGTCCGTGGTGTGGTAGGACTGTAGAA-------AAGGG   1405

Qy  1484  AACTGAACATTCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTT   1543
```

```
Db   127 CGAGGAAATGGGCTCAGGGGACTATGACTTCCATGAAGGAACCCTGTTCCGTGAAGAAAA 186
         |||||||||||||||||||||||||||||| ||||||      ||||      ||||||||
Qy   224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCATCTCTTGACTGGCAT 283
         || |||||||||||  ||||||||||||||||||||||||||||||  |||||||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGGATTGGTGATCGATCGGTCATGGGTCATGGGTTACCAGAAGAGCTAAGGAGCATGAC 343
         |||||||||||||||||| |||| ||||||||||||        |||||||||||||||||||||||||
Db   247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCAGTCCTGTCAGTGGCTCAGTGCTGACCTCCTTTGTCATCACACTCCCCTT 403
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db   307 GGACAAGTACAGGCTGCACCTGCAGTCGTCAGTGGCCGACCTCCTTTGTCATCACGCGTTCCCTT 366

Qy   404 CTGGGCAGTTGATGCCATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGTCCA 463
         |||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db   367 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTCTCACAGCACGGTTCTCATCCTGGCCTTCATCTGGCCTCTGGA 523
         |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   427 TGTCATCTACACAGTCAACCTCTACACAGCAGTGTCTTCATCCTCATCCTGGCCTTCATCAGTTGGA 486

Qy   524 CCGGTACCTCGCCATTGTCCAGCGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGA 583
         ||| ||||||| ||||| ||||||||| ||||||||||||||||||||||||||| |||||
Db   487 CCGCTACCTGCCCATCGTCCGCCACCAACAGTCAGAGGCCAAGAAGCTGTTGGCTGA 546

Qy   584 AAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGACTATACTGACTTCAT 643
         |||||||||||||||||||  |||||| |||||| |||| ||||||| |||||||||
Db   547 AAAGGTGGTCTATGTTGGGCGTCTGGATCATCGTGCCTGCTGACTATTCCCGACTTCAT 606

Qy   644 CTTTGCCACGTCAGCCAGGGACATCAGTGGGGGGAAATCGTGGCTGCTGA--------- 703
         |||||||                    ||||||||||||||||||||||
Db   607 CCGCTACCTGCCCATCGTCCGCCACCAACAGTCAG------TCAGTGAGGCAGATGATGACAGATATCGTGACCG 651

Qy   704 CCCTTACCCCGATAGCCGTGTGGATGTGGTGTTCAATTCCAGCATATAATGGTGGGTCT 763
         |||||||||||||||||||||||||||||||||||| |||||| || ||||||| |||||
Db   652 CTTCTACCCCCAATGACTTGTGTGGGTGTGGTGTTCAGTTGTTCAGCACCATGCATGGTTGGCCT 711

Qy   764 CATCCTGCCGGCATCGTCATCCTCCTGTTACTGCTGATCATCATCTCTAAGCTGTCACA 823
         ||||||||| |||||||||||||||||||| |||| |||||| ||||||||||||||||
Db   712 TATCCTGCCTGGTATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCCAAGCTGTCACA 771

Qy   824 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGTGCATCATCCTCAGCTTT 883
         ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||| |
Db   772 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACCACCACAGTCATCCTCATCCTGGCTTT 831

Qy   884 CTTTGCTGCTGGCTGCCATTATTGTGGGGATCGACGACTCCTTCATCCTTTTGGG 943
         ||||| ||||| |||| ||| |||||||| |||| |||| || |||||| |||||
Db   832 CTTCGCTGTTGGCTGCCTTCACAGTTGCTACTACATTGGGATCGACGACTCCTTCATCCTCTGA 891

Qy   944 AGTCATCAAGCAAGGAGATGGACTTCGAGAGACGCATTGTGCACAAGTGGATCTCCATCACAGA 1003
         |||||||||||||||| || |||  |||| ||||  ||||| ||||||||||||||| ||||||
Db   892 AATCATCAAGCAAGGGTGTGAGTTTGAGAACAACTGTGCACAAGTGGATTCCATCACCGA 951

Qy  1004 GGCCCTCGCCTTCTTCCAGCGTGCCTGGCCCATCCTCATGAGCGAGAGGGTCCGGGGCCAA 1063
         |||| |||| ||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   952 GGCCCTAGCTTTCTTCCACTGTGTCTGAACCCATCCTCTATGCTTTCCTGTTCCTTGGAGCCAA 1011

Qy  1064 GTTCAAAAGCTCTGCCCAGATGCAGCTCAACTCAACTCAGAGCGAGAGGTCCAGCCTCAAGAT 1123
         ||||||||| |||||||| ||| |||||||||| ||| || |||||||||||||||||||||
Db  1012 ATTTAAAACCTCTGCCCAGACGCGCACTCACCTCTGTCCGTGTCTCCACGGACTCAGAATCCTCCAG 1071

Qy  1124 CCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTGCTCCGGTAACCACCACGGCTG 1183
         |||||||||||||||||||| |||| ||||||| ||||  ||| ||| | || |||
Db  1072 CCTCTCCAAAGGAAAGCGAGGTGGACATTCCACTGCATTCATCTGTTCCACTGAGTCTCTTCAAG 1131

Qy  1184 TTTTCACTTCCAGCTAACCCTTATGCAAAGACTTATATATATACACTTCTGATAACTACTC 1243
         ||||||||| ||| ||||||| ||||| ||||||||||||||||||      |||||||||
Db  1132 TTTTCACTTCCAGCTAACCCTTATGCAAAGACTTATATATATACACTTCT------TTTATACGATAAATA 1183

Qy  1244 GAGAACCAGCGCGGTTACCCATGGA----GGGGATCCAGTATATAACACTTCTACTC 1303
         ||||    | |||||||  |||||||||||||||||||  ||||||| || |||
Db  1184 GAACTTTTTTAGTTACCACATTTTTCAGATATAAAGACTGACCAATATTGTACAGTTTT 1243
```

RESULT 7

AAI65467

Query Match   50.8%; Score 952.8; DB 4; Length 1679;
Best Local Similarity 80.2%; Pred. No. 4.1e-174;
Matches 1210; Conservative   0; Mismatches 252; Indels   46; Gaps   6;

```
Qy    44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
         || |||||||||||||| || || |||  |||||||||  |||||||||||||||| |||
Db    12 GCGGCAGCAGGTAGCAGAAGTGACGCGGAGGGCCCGAGGTGCTCCAGTAGCCACCATCTG 71

Qy   104 TAGAGCGAGTGTTGCCATGGAACCATGGA----GGGGATCCAGTATATAACACTTCTACTC 163
         |||||| |||||||||||| |||||| ||||||||||||||  |||||| || |||
Db    72 GAGAACCAGCGCGGTTACCCATGGA-----GGGGATCCAGTATATAACACTTCTACTC 126

Qy   164 TGAAGAAGTGGGGTCTGGAGACTATGACTTCCAACAAGGAACCCTGTTCCGGGATGAAAA 223
         || |||| ||||||| |||| ||||||||||||
Db          TGAAGAAGTGGGGTCTGGAGACTATGACTTCCAACAAGGAACCCTGTTCCGGGATGAAAA
```

Db 1406 AACTGAACATTCCAGAGAGCGGGGTGAGTGAATCACGCACGTAAAGCTAGAAATGATCCCCAGCTGTT 1465
Qy 1544 GCTGCATA 1551
Db 1466 TATGCATA 1473

Qy   1304 TTTTTTTTTTTAATGACTGTGGGAGTTATGTTCCTCAGTTTTTGTGAGGTTTGACT 1363
           ||| || |||| ||| |||||||||||||||||||
Db   1244 TATTGCTTGTTGGATTTTGT-----CTTGTGTTCTTTAGTTTTTGTGAAGTTTAATT 1297

Qy   1364 TAATTTATATAAAATATTGTTTTTTGTTTTGTTTCATGTGAATGAGCGTTCAGGCAGGACCT 1423
           |||||| ||||||| |||| ||||||||| || || |||||||||||||||
Db   1298 GACTTATTTATA---TAAATTTTTTTTGTTTCATTATGATGTGTGTCCAGGCAGGACCT 1353

Qy   1424 GTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTAGGACTGTAGAACTGTAGAGGAGA 1483
           |||||||||||||||||| |||| |||||| ||||||||||||| |||||
Db   1354 GTGGCCAAGTTCTTAGTTGCTGTATGTCTCGTGGTAGGACTGTAGAA-------AAGGG 1405

Qy   1484 AACTGAACATTCCAGAAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTT 1543
           |||||||||||||||| ||||| || ||||| ||||||||||||| ||||||||||
Db   1406 AACTGAACATTCCAGAGCCGTGTAGTGCGAATCACGTAAAGCTAGAAATGATCCCCAGCTGTT 1465

Qy   1544 GCTGCATA 1551
           | |||||||
Db   1466 TATGCATA 1473

RESULT 8
ABSS3992
ID   ABSS3992 standard; cDNA; 1679 BP.
XX
AC   ABSS3992;
XX
DT   02-DEC-2002 (first entry)
XX
DE   DNA encoding human CXC chemokine receptor 4 (CXCR4).
XX
KW   CXC chemokine receptor 4; CXCR4; chemotaxis induction; SDF1a;
KW   stromal cell-derived factor 1alpha; angiogenesis; atherosclerosis;
KW   restenosis; ischaemic stroke; spinal cord injury; infection; ulcer;
KW   human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;
KW   AIDS; pain; cancer; benign prostatic hypertrophy; diabetes; obesity;
KW   anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW   hypotension; hypertension; urinary retention; osteoporosis; stroke;
KW   anginal pectoris; myocardial infarction; benign prostatic hypertrophy;
KW   migraine; vomiting; psychotic disorder; neurological disorder; anxiety;
KW   schizophrenia; dyskinesia; Huntingdon's disease; restenosis;
KW   inflammatory disease; rheumatoid arthritis; diabetic retinopathy;
KW   inflammatory bowel disease; atherosclerosis; Alzheimer's disease;
KW   congestive heart failure; cardiac remodeling; angiogenic diseases;
KW   solid tumour; Kaposi Sarcoma; human; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   89..1147
FT   /tag= a
FT   /product= "CXCR4"
FT   /note= "CXC chemokine receptor 4"
XX
PN   US2002107195-A1.
XX
PD   08-AUG-2002.
XX
PF   17-SEP-2001; 2001US-00953692.
XX
PR   21-JUL-1998; 98US-0093596P.
PR   21-JUL-1999; 99US-0358624.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI   Gupta SK;
XX
DR   WPI; 2002-697879/75.
DR   P-PSDB; ABG32977.
XX
PT   Inducing chemotaxis of endothelial cells, useful for regulating
PT   angiogenesis, e.g. for treating cancer, comprises treatment with stromal

PT   cell-derived factor 1 alpha.
XX
PS   Disclosure; Fig 1; 26pp; English.
XX
CC   The invention describes a method of inducing chemotaxis of endothelial
CC   cells by treatment with stromal cell-derived factor 1alpha (SDF1a). The
CC   method is used for stimulating EC chemotaxis, and thus angiogenesis, and
CC   is used for treating atherosclerosis, restenosis, ischaemic stroke and
CC   spinal cord injury. Inhibition of this process is useful in treatment and
CC   prevention of a very wide range of diseases, such as, infection (by
CC   bacteria, fungi, protozoa or viruses such as human immunodeficiency virus
CC   (HIV) and acquired immunodeficiency syndrome (AIDS)), pain, cancer and
CC   benign prostatic hypertrophy, diabetes, obesity, anorexia, bulimia,
CC   asthma, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, anginal pectoris,
CC   myocardial infarction, stroke, ulcers, benign prostatic hypertrophy,
CC   migraine, vomiting, psychotic and neurological disorders (e.g. anxiety,
CC   schizophrenia) and dyskinesias (e.g. Huntingdon's disease), inflammatory
CC   diseases, rheumatoid arthritis, diabetic retinopathy, inflammatory bowel
CC   disease, atherosclerosis, restenosis, Alzheimer's disease, congestive
CC   heart failure, cardiac remodeling, angiogenic diseases, solid tumours,
CC   and Kaposi Sarcoma. This sequence encodes the human CXC chemokine
CC   receptor 4 (CXCR4) polynucleotide
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   50.8%; Score 952.8; DB 6; Length 1679;
Best Local Similarity   80.2%; Pred. No. 4.1e-174;
Matches 1210; Conservative   0; Mismatches 252; Indels   46; Gaps   6;

Qy   44 GCAGGTGCAGGTAGCAGTTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
           || ||| |||||||||| || ||| ||| | ||| |||||| || |||||||| |||
Db   12 GCGGCAGCAGGTAGCAAAGTGACAGCCGCAGGGCCTGAGTGCTGCTCCAGTAGCACCATCTG   71

Qy   104 TAGAGGCAGTGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
           || | ||| ||||||||| |||| |||||||||||||||||||||| ||||||||||
Db   72 GAGAACCAGCGGTTACCATGGA----GGGGATCAGTCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
           ||| || |||| ||||| ||||||||||| ||||||||||||| |||||| ||| |||
Db   127 CGAGGAAATGGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy   224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTGACTGGCAT 283
           |||| |||||| | ||||||||||||||||||||||||||||||||||| |||||||||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGTAAGGAGCATGAC 343
           | | ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   247 TGTGGGCAATGAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAAACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGTCAGTGCTGACCTCCCTTGTCATCACACTCCCCTT 403
           |||||||||| |||||||||||||||||||| ||| |||||||||||| || ||||
Db   307 GGACAAGTACAGGCTGCACCTGTCAGTGCTGGCCGACCTCCTTGTCATCACGCTTCCCTT 366

Qy   404 CTTGGGCAGTTGATGCCATGGCTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
           |||||||||||||||| |||| |||||||||||||| || | |||||| |||||
Db   367 CTTGGGCAGTTGATGCCGTGGCAAGTGGTACTTTGGGAACTTCATGCAAGGCAGTCCA 426

Qy   464 TATCATCACACTGTCAACCTCTACAGCACGGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
           | ||||| |||| ||||||||||||||| ||| || ||||||||||||||||| |||||
Db   427 TGTCATCTCACACAGTCAACCTCTACAGCAGTGTCCTATCCTGGCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
           ||| ||||| |||| |||||||||||||||||||| |||||||||| || ||||||||
Db   487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGGAGCTGTTGGCTGA 546

Qy   584 AAAGGCAGTCTATGTGGGGCTCTGGATCCCAGCCCTCCTGCTGACTATACCCTGACTTCAT 643
           ||||| |||||||| |||| | ||||||||| || ||||||| | ||| ||| ||||||||
Db   547 AAAGGTGGTCTATGTTGGGCGTCTGGATCCCGCTCTGGATCCTGCTGTCGTCTATTCCCGACTTTCAT 606

Qy   644 CTTTGCCGCAGTCAGCCAGGCGGACATCAGTCAGGGGGATGACAGGTACACATCTGTGACCG 703
           |||||||| | | |||||||||||||| || ||||||||||||
Db   607 CTTTGCCAACG-----------TCAGTGAGCAGATGACAGATATATCTGTGACCG 651

KW   antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;
KW   hypoxia; tumourigenesis; angiogenesis; apoptosis; cancer;
KW   ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
KW   preeclapsmia; atherosclerosis; inflammatory condition; wound healing;
KW   inflammation; erythropoiesis; hair loss; human; gene; ds..
OS   Homo sapiens.
XX
PN   WO20024665-A2.
XX
PD   13-JUN-2002.
XX
PF   10-DEC-2001; 2001WO-GB005458.
XX
PR   08-DEC-2000; 2000GB-00030076.
PR   08-FEB-2001; 2001GB-00003156.
PR   25-OCT-2001; 2001GB-00025666.
XX
PA   (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI   White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
PI   Rayner WN;
XX
DR   WPI; 2002-627238/67.
XX
PT   Identifying a gene involved in disease for treating hypoxia-regulated
PT   conditions, comprises comparing the transcriptome/proteome of two cell
PT   types under different conditions and identifying a differentially
PT   regulated gene.
XX
PS   Claim 37; Page 378; 538pp; English.
XX
CC   The present invention relates to methods for identifying genes and
CC   proteins that are implicated in a specific disease or physiological
CC   condition. The method comprises comparing the transcriptome/proteome of a
CC   specialised cell type implicated in a disease or condition with that of a
CC   second specialised cell type, under two experimental conditions, and
CC   identifying a gene that is differentially regulated in the two
CC   specialised cell types under experimental conditions. ABV77873-ABV78116
CC   and ABP65061-ABP65257 were identified using the methods of the invention.
CC   The coding sequences and proteins are useful for treating a disease in a
CC   patient, for manufacture of a medicament for treating hypoxia-regulated
CC   conditions, and for regulating tumourigenesis, angiogenesis, apoptosis,
CC   biological response to hypoxia conditions, or hypoxic-associated
CC   pathology in a patient. The coding sequences and proteins are also useful
CC   for monitoring the therapeutic treatment of a disease or physiological
CC   condition, such as cancer, ischaemic conditions, reperfusion injury,
CC   retinopathy, neonatal stress, preeclapmia, atherosclerosis, inflammatory
CC   conditions, wound healing, inflammation, erythropoiesis or hair loss
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   50.8%;   Score 952.8;   DB 6;   Length 1679;
Best Local Similarity   80.2%;
Matches 1210;   Conservative   0;   Mismatches   0;   Indels   46;   Gaps   6;
Pred. No. 4.1e-174;

```
Qy     44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCGTTGGTGCTCCGGTAACCACCAGGGCTG 103
          |||||||||||||| ||||||||||||| |||||||||||| ||||||||||||||||
Db     12 GCGGCAGCAGGTAGCAAAGTGACACGCCGAGGGCCTGAGTGCTGCTCCAGTGCCGCATCTG  71

Qy    104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
          ||||| ||||||||||||||||||||||||||||||     |||||||||||||||||||
Db     72 GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTAGTATATACACTTCAGATAACTACAC 126

Qy    164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
          |||||||||||| ||||||||||||||||||| |||||||||||||||||||||||||||
Db    127 CGAGGAAATGGGCCTCAGGGGACTATGACTCCAGAGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy    224 CGTCCATTCAATAGGATCTTCCTGCCACCATCACTTCATCATCTTCTTGACTGGCAT 283
          |||||||| |||| ||  ||||||||||||||||||||||||||||||||||| ||||
Db    187 TGCTAATTTCAATAAAATCTTCCTGCCACCATCTCCTCACTCACTCTTCTTAACTGGCAT 246

Qy    284 AGTCGGCAATGAATGGATTGGTGATCCTGGATCCTGGTCATCGTGGTTACCAGAAGAAGAAGTAAGGAGCATGAC 343
```

```
Db  1298 GACTTATTTATA---TAAATTTTTTTGTTCATATTGATGTGTGTCTAGGCAGGACCT 1353
              ||||||||||        || ||||||||||||||||| || |||||||||||||||||
Qy  1424 GTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAGGAGAGGAAGA 1483
         ||||||||||||||||| |||| |||| |||||| |||||||| ||||||
Db  1354 GTGGCCAAGTTCTTAGTTGCTGTGTATGTCTCGTGGTAGGACTGTAGA------AAGGG 1405

Qy  1484 AACTGAACATTCCAGAATGTGTGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTT 1543
         |||||||||||||||| | ||| || ||| ||||||||||| |||| ||||| |||||||
Db  1406 AACTGAACATTCCAGAGCGTGAGTAGTGAATCACGTAAAGCTAGAAATGATCCCCAGCTGTT 1465

Qy  1544 GCTGCATA 1551
         || |||||
Db  1466 TATGCATA 1473
```

RESULT 10
ABX08779
ID   ABX08779 standard; cDNA; 1679 BP.
XX
AC   ABX08779;
XX
DT   21-JAN-2003 (first entry)
XX
DE   Angiogenesis-associated human polynucleotide sequence #41.
XX
KW   Human; angiogenesis-associated transcript; angiogenesis;
KW   angiogenesis-associated disease; cancer; cytostatic; gene therapy; gene;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200279492-A2.
XX
PD   10-OCT-2002.
XX
PF   14-FEB-2002; 2002WO-US004915.
XX
PR   14-FEB-2001; 2001US-00784356.
PR   22-FEB-2001; 2001US-00791390.
PR   19-APR-2001; 2001US-0285475P.
PR   03-AUG-2001; 2001US-0310025P.
PR   13-NOV-2001; 2001US-0350666P.
PR   29-NOV-2001; 2001US-0334244P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Murray R, Glynne R, Watson SR, Aziz N;
XX
DR   WPI; 2003-040681/03.
DR   P-PSDB; ABU03495.
XX
PT   Detecting angiogenesis-associated transcript in a cell for diagnosing and
PT   treating cancer by contacting a sample with a polynucleotide that
PT   exhibits changes in expression level as a function of time in tissue
PT   undergoing angiogenesis.
XX
PS   Example 2; Page 217-218; 291pp; English.
XX
CC   The present invention relates to methods and compositions for detecting
CC   an angiogenesis-associated transcript in a cell in a patient. The method
CC   involves contacting a biological sample from the patient with a
CC   polynucleotide that selectively hybridises to a sequence at least 80%
CC   identical to any of the angiogenesis-associated human polynucleotide
CC   sequences given in the specification. These angiogenesis-associated
CC   polynucleotide sequences comprise genes that exhibit changes in
CC   expression levels as a function of time in tissue undergoing
CC   angiogenesis. The method and the polynucleotide sequences of the
CC   invention are useful for diagnosing and treating angiogenesis and
CC   angiogenesis-associated diseases e.g. cancer. The polynucleotide
CC   sequences are also useful in the gene therapy of such disorders. The
CC   angiogenesis-associated proteins encoded by the polynucleotide sequences
CC   are useful as a vaccine for therapeutic and prophylactic immunisation.
CC   ABX08739-ABX08853 represent angiogenesis-associated polynucleotide
```

```
         247 TGTGGGCCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAGAAACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGCTCAGTGGCTGACCTCCTCTTTGTCATCACGCTTCCCYT 403
Db   307 GGACAAGTACAGGCTGCACCTGCTCAGTGGCCGACCTCCTCTTTGTCATCACGCTTCCCYT 366

Qy   404 CTGGGCAGTTGATGCCATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGTCCA 463
Db   367 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTTCTACAGCACGCGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
Db   427 TGTCATCTACACAGTCAACCCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTCGGCCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
Db   487 CCGCTACCTGGCCATCGTCCACCGCGCCACCAACAGTCAGAGGCCAAGGAAGCTGTGTTGGCTGA 546

Qy   584 AAAGGCAGTTCATGTTGGGCGTTCTGGATCCCGACCCTCCTCCTGACTATACCTGACTTCAT 643
Db   547 AAAGGTGTCTCTATGTTGGCGTCGGGATCCCTGCCCTCCTGCTGACATTCCCGACTTCAT 606

Qy   644 CTTTGGCCGACGTCAGCAGGGGACATCAGTCAGGGGATCACAGGTACATCTGTGACCG 703
Db   607 CTTTGGCCAACG----------TCAGTGAGGCAGATGACAGATATCTGTGACCG 651

Qy   704 CCTTTACCCCGATAGCCGTGTCGTGGTGTTCAATTCCAGCATATAAATGGTCGGTCT 763
Db   652 CTTCTACCCCCAATGACTTGTCGGGTGGTGTGTGTTCAGCACACATCATGGTGGCCT 711

Qy   764 CATCCCTGCCCGGCATCGTCATCCTCTCCGTGTACTGCATCATCTCTAAGCTGTCACA 823
Db   712 TATCCTGCCTGGTATTGTCATCCTGCTATTGCATTTCATCTCCAAGCTGTCACA 771

Qy   824 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACAGCAGTCATCCTCATCCTAGCTTT 883
Db   772 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACAGCAGTCATCCTCATCCTGCCTTT 831

Qy   884 CTTTGCTGCTGGCTGGCCATATTATGTGGGGATCAGCAGTCAGCATTCACCTTTTGGG 943
Db   832 CTTCGCCCTGTTGGCTGCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTCCTGGA 891

Qy   944 AGTCATCAAGCAAGGATGTGACTTCGAGAGAGCATTGTGCACAAGTGGATCTTCCATCACAGA 1003
Db   892 AATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGA 951

Qy  1004 GGCCTGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTCCTGGGGCCAA 1063
Db   952 GGCCCTAGCTTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTTCTTCCTTGGGAGCCAA 1011

Qy  1064 GTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATGAGACAGAGGCTCCAGCCTCAAGAT 1123
Db  1012 ATTTAAAAACCTCTGCCCAGCACGCCACTCACCTGTGTGAGCAGAGGGTCCAGCCTCAAGAT 1071

Qy  1124 CCTTTCCAAAGGAAAGGCGGGGTGGACACTCTTCATCATCGTTTCCACTGTGTTCCACTGAGTCTGAGTCTTCAAG 1183
Db  1072 CCTCTCCAAAGGAAAGGCGAGGTGGACATTCATCATCGTTTCCACTGAGTCTGAGTCTTCAAG 1131

Qy  1184 TTTTCACTCCAGCTAACCCTTATGCAAAGACTTATATAATATATAATATATGATAAA 1243
Db  1132 TTTTCACTCCAGCTAACAGATGTAAAAGACTTTTT------TTTATACGATAAATA 1183

Qy  1244 GAACTTTTTATGTTACACATTTTCCAGATATAAGAGACTAGACTGTCTGTACAGTTTT 1303
Db  1184 ACTTTTTTTTAAGTTACACATTTTCAGATATAAAAGACTGACCCAATATTGTACAGTTTTT 1243

Qy  1304 TTTTTTTTTTAATTGACTTGTGGGAGTTTATGTTCCCTAGTTTTTGTGAGGTTTGACT 1363
Db  1244 TATTGCTTGTTGGGATTTTTGT------CTTGTGTTCTTTTAGTTTTTAGTTTTTGTGAAGTTTAATT 1297

Qy  1364 TAATTTATATAAAATATTGTTTTTTGTTTGTTTCATGTGAATGAGCGGTCTCAGGCAGGACCT 1423
```

```
CC  sequences
XX
SQ  Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match    50.8%;  Score 952.8;  DB 7;  Length 1679;
Best Local Similarity  80.2%;  Pred. No. 4.1e-174;
Matches 1210;  Conservative  0;  Mismatches 252;  Indels  46;  Gaps  6;

Qy    44  GCAGGTGCAGTAGCAGTGACCCTGAGCGTTGGTGTGTCCGGTAACCACCACGGCTG   103
          ||||||| |||||| |||| |||||||||||||||||| |||||||||||||||||
Db    12  GCGGCAGCAGGTGCAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG    71

Qy   104  TAGAGCCAGTGTTGCCATGGACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC   163
Db    72  GAGAACCAGCCGGTTACCATGGA------GGGGATCAGTATATACACTTCAGATAACTACAC   126

Qy   164  TGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCCGGATGAAAA   223
Db   127  CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTCCGTGAAGAAAA   186

Qy   224  CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTGACTGGCAT   283
Db   187  TGCTAATTTCAATAAAATCTTCTGCGCCACCATCTACTCCATCATCTTCTAACTGGCAT   246

Qy   284  AGTCGGCAATGGATTGGTGATCCTGGTTCATGGGTTACCAGAGAGCTAAGGAGCATGAC   343
Db   247  TGTGGGCAATGATTGGTCATCCTGGTCATGGGTTACCAGAGAAGAACTGAAGAAGCATGAC   306

Qy   344  GGACAAGTACCCGGCTGCCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTT   403
Db   307  GGACAAGTACAGCGGCTGCCACCTGTCAGTGGCCGACACCTCCTCTTGTCATCACGCTTCCCTT   366

Qy   404  CTGGGCAGTTGATGGCCATGGCTGACTGTACTTTGGGAAATTTGTGTAAGGCTGTCCA   463
Db   367  CTGGGCAGTTGATGGCCATGGCTGACTGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA   426

Qy   464  TATCATCTACACGTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGA   523
Db   427  TGTCATCTACACAGTGCTAACCTCTACAGCAGTGTCCTCATCAGCCTTCATCAGCCTGGA   486

Qy   524  CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA   583
Db   487  CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA   546

Qy   584  AAAGGCAGTTCTATGTGGGCGTCTGGATCCCTCCTGAGATCCAGCCCTTCCTGACTTCAT   643
Db   547  AAAGGTGGTTCTATGTTGGCGTCTGGATCCCTGCCCTCCTGTGCCCTGACTATTCCCGACTTCAT   606

Qy   644  CTTTGCCACGTCAGCCACCTCAGGGGACATCAGTCAGGGGGATGACAGCAGTTACATCTGTGACCG   703
Db   607  CTTTTGCCAACG-----------TCAGTGAGGCAGGCAGATGACAGATATATCTGTGACCG   651

Qy   704  CCTTTACCCGGATAGCCTTGGATGGTGTGTTCAATTCCAGGCATATAATGGTGGGTCT   763
Db   652  CTTCTACCCCAATGACTTTGGGGTGGTTGTTGCTGTTCCAGTTTCAGCACACTATCTGGCCT   711

Qy   764  CATCCTGCCCGCCGGATCCGTCATCCTCCTGTACTACTGACCGTCATCCTCTAAGCTGTCACA   823
Db   712  TATTCCTGCCTGGATTGTCATCCTCGTCTTGCTATTGCATTATCATTCTCCAGCTGTCACA   771

Qy   824  CTTCCAAGGGCCACCAGAAGCGCAAGAGCCCTCAAGACGACCAGTCATCCCTCATCCCTAGCTTT   883
Db   772  CTTCCAAGGGCCACCAGAGCGCAAGGCCCTCAAGACCAGTCATCATCCTCATCCTGGCTTT   831

Qy   884  CTTTGCCTGCTGGCTGCCATATTAATGTGGGATCAGCATCGACTCCTTTCATCCTTTTGGG   943
Db   832  CTTCCGCCTGTTGGCTGCTGCCTTACTACAATTGGGATCAGCATCGACTCCTTTCATCCTCCTGGA   891

Qy   944  AGTCATCAAGCAAGGATGGACTTCGAGACATTGTGCACAAGTGGATCTCCATCACAGA   1003
Db   892  AATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTGCCACAAGTGGATTTCCATCACCGA   951

Qy  1004  GGGCCCTGGCCTTCTTCCCACTGTTGCCTGAACCCCATCCTCTATGCTTCCTCGGGGCCAA   1063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   952  GGCCCTAGCTTCTTCTTCCACTGTTGCTGAACCCATCCTCTATGCTTCCTTGGAGCCAA   1011

Qy  1064  GTTCAAAAGCTCTGCCCAGCATGACCATGAGCAATCCAACTCCATGAGCAGAGGCTCCAGCCTCAAGAT   1123
Db  1012  ATTTAAAACCTGCCCAGCACGCAGCCACTCACCTGTGAGCAGCAGGGTCCAGCCTCAAGAT   1071

Qy  1124  CCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGAGTCAGAATCCTCCAG   1183
Db  1072  CCTCTCCAAAGGAAAGCGAGGTGGACATTCATCGTTTCCACTGAGTCTGAGTCTTCAAG   1131

Qy  1184  TTTCACTCCAGTCAACCCTTATGCAAGACTTATATATATATATATATATATGATAAA   1243
Db  1132  TTTCACTCCAGCTAACACAGATGTAAAGACTTTT------TTTATACGATAAATA   1183

Qy  1244  GAACTTTTTATGTTACACATTTCCAGATATAGAGACTGACCAGTCTGTACAGTTTT   1303
Db  1184  ACTTTTTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTT   1243

Qy  1304  TTTTTTTTTTTTAATTGACTGTGGGAGTTTTATGTTCCTAGTTTTTGTGAGGTTTGACT   1363
Db  1244  TATTGCTTGTGTGGATTTTGT------CTTGTGTTTTCTTTAGTTTTTGTGAAGTTTTAATT   1297

Qy  1364  TAATTTATATAAAATATTGTTTTTTTGTTCATGTGAATGAGCGTCTAGGCAGGACCT   1423
Db  1298  GACTTATTTATA-----TAAATTTTTTTTGTTCATATTGATGTGTCTTAGGGCAGGGACCT   1353

Qy  1424  GTGGCCAAGTCTTAGTAGCTGTTTATCTGTGTGGAGCTGTGAGAACTGTAGAGGAAGA   1483
Db  1354  GTGGGCCAAGTTCTTAGTTGCTTAGTTGTCTCGTGGGTAGGACTGTAGAA------AAGGG   1405

Qy  1484  AACTGAACATTCCAGAAATGTGTGTAAATTGAATAAAGCTAGCCGTGATCCTCCAGCTGTT   1543
Db  1406  AACTGAACATTCCAGAGCGTGTAGTGAATCCGTAAAGCTAAGAAATGATCCCCAGCTGTT   1465

Qy  1544  GCTGCATA   1551
          || |||||
Db  1466  TATGCATA   1473
```

```
RESULT 11
ABX74454
ID   ABX74454 standard; cDNA; 1679 BP.
XX
AC   ABX74454;
XX
DT   21-MAR-2003  (first entry)
XX
DE   Human cDNA sequence #26 up-regulated in CC-RCC patients.
XX
KW   Human; microarray; solid surface; immobilised probe; CC-RCC;
KW   differential expression profile; aggressive CC-RCC tumour type;
KW   non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
KW   gene expression profiling; tumour tissue; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200279411-A2.
XX
PD   10-OCT-2002.
XX
PF   29-MAR-2002; 2002WO-US009576.
XX
PR   29-MAR-2001; 2001US-0279411P.
XX
PA   (VAND-) VAN ANDEL INST.
XX
PI   Haab B, Rhodes D, Teh BT, Takashi M;
XX
DR   WPI; 2003-040679/03.
XX
PT   New microarray, comprising a matrix of cDNA probe from a set of probes
PT   immobilised to a solid surface in predetermined order, useful in the
```

```
PT   prognosis of patients with clear cell renal carcinoma.
XX
PS   Claim 35; Page 132-133; 179pp; English.
XX
CC   The present invention relates to a microarray comprising a matrix of at
CC   least one cDNA probe from a set of probes immobilised to a solid surface
CC   in a predetermined order, where a row of pixels corresponds to replicates
CC   of one distinct probe from the set. The probes are complementary to
CC   nucleic acid sequences that are expressed differentially in aggressive as
CC   compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
CC   and that hybridise to the probes under high stringency conditions. The
CC   microarray is useful for the prognosis of patients with CC-RCC, wherein
CC   aggressive and non-aggressive CC-RCC tumour types are characterised by
CC   differential expression profiles of genes that hybridise with one or more
CC   probes immobilised on the microarray. The arrays are useful for gene
CC   expression profiling of tumour and normal tissues. The present sequence
CC   represents a human cDNA sequence up-regulated in CC-RCC patients
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match     50.8%;  Score 952.8;  DB 7;  Length 1679;
Best Local Similarity   80.2%;  Pred. No. 4.1e-174;
Matches 1210;  Conservative    0;  Mismatches  252;  Indels  46;  Gaps    6;
```

```
Qy      44 GCAGGTGCAGTAGCAGTGACCCTGAGCGTTGGTGTCCGTAACCACCACGGCTG 103
           ||||||| |||||||||||| |||  | |||||||| | |||||| |||||||
Db      12 GCGGGAGCAGTAGCAAAGTGACGGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCACCCGCATCTG 71

Qy     104 TAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
           |||||||||||||||| |||||||||||||||| ||||||||||||||| |||||||||||
Db      72 GAGAACCAGGGGTTACCATGA------GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy     164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAGGAGACCCTGCTTCCGGATGAAAA 223
           ||| |||||||| || |||||||||||||||||||||||| |||||| ||||||||||
Db     127 CGAGGAAATGGGCTCAGGGGACTATGACTCCAATGAAGAACCCTGTTTCCGTGAAGAAA 186

Qy     224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTGACTGGCAT 283
           ||||||||||||||  |||||||||||||||||||||| ||||||||||||||||||
Db     187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCACTCCATCATCTTCTAACTGGCAT 246

Qy     284 AGTCGGCAAATGGATTGGTGTCATGGTCTGTGTACCAGAAGAAGCTAAGGAGCATGAC 343
           ||||||||||||||||||||||||||||||||||||| ||||||||||| ||||||||
Db     247 TGTGGGCAATGGATTGGTGTCATGGTCTGTGTACCAGAGAAGAAACTGAGAAGCATGAC 306

Qy     344 GGACAGTACCGGCTGCACCCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCT 403
           |||||||||| | | |||||||||||||||||||||||| |||||||||||| ||||||
Db     307 GGACAGTACAGGCTGCACCCTGTCAGTGGCCGACCCCTCTTGTCATCACGCCTTCCCT 366

Qy     404 CTTGGGCAGTGATGCCATGGCTGACTGGGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
           |||||||||||||||||||||||||||||||| |||||||||||| |||||||||||||||
Db     367 CTTGGGCAGTTGATGCCGCTGGCAAACTGGTACTTTGGGAAACTTCCTATGCCAAGGCAGTCCA 426

Qy     464 TATCATCTACACTGTCAACCCTACAGCACGCGTTCTCATCCTCGGCCTTCCATCCAGCAGCCTGGA 523
           |||||||||||||| |||||||||| |||||| ||| ||||||||| ||||||||||||||||||
Db     427 TGTCATCTACACAGTCAACCCTCTACACGCGTCCTCCTCCCTCGGCCTTCCATCCAGCAGTCTCGA 486

Qy     524 CCGGTACCTCGCCATTGTCCAGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGA 583
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     487 CCGCTACCTCGCCATGGCCATGCTCCACGCGCCACCAACAGTCAAGGAGAAGGTGTTGGCTGA 546

Qy     584 AAAGGCAGTCTATGTGGGCGTCTGATCCCAGCGCCCTCCTCCGACTATACCTGACTTCAT 643
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     547 AAAGGTGGTCTATGTGGGCGTCTGGATCCCGCCTCCTGCCCTGCTGACTTCAT 606

Qy     644 CTTTGCGGACGCTGCAGCCAGGCGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCG 703
           |||||||||||||||||||||||        |||||||||||||||||||||||||||||||
Db     607 CTTTGCCAACG------------TCAGTGAGGCAGATGACAGATATCTGTGACCG 651

Qy     704 CCTTTACCCCGATAGCCTGTCGATGGTGTGTTCAATTCCAGCATAAATGGTGGGTCT 763
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     652 CTTCTACCCCAATGACTTGTGGGTGGTTTGTGTTCCAGTTTCAGCACATCATGGTGGCCT 711

Qy     764 CATCCTGCCCGGCCATCGTCATCCTCCTGTTACTGCCATCATCATCATCCTTCAAGCTGCACA 823
```

```
Db     712 TATCCTGCTGGTATTGTCATCCTCGTCCTGCTATTGCATTATCATCTTCCAAGTGTCACA 771
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy     824 CTCCAAGGCGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGCAGTCATCCTCATCCTAGCTTT 883
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     772 CTCCAAGGCCACCAGAAGCGCAAGGCCCTCAAGACCAGTCATCCTCATCCTCGGCTTT 831

Qy     884 CTTTGCCTGCTGGTGCCATATTATGTGGGATCAGCATCAGACATCGACTCCTTCATCCTTTTGGG 943
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     832 CTTCGGCCTGTTGGCTGCTTACTACATTGGGATCAGCATCAGACATCGACTCCTTCATCCTCCTGA 891

Qy     944 AGTCATCAAGCAAGGATGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGA 1003
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     892 AATCATCAAGCAAGCAAGGGGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGA 951

Qy    1004 GGCCTGCCTTCTTCCACTGTTGCCTGAACCCATCCTCATGCCTTCCTCGGGGCCAA 1063
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     952 GGCCTAGCTTCTTCTTCCACTGTTGCTGAACCCATCCTCATGCTTTCCTTGGAGCCAA 1011

Qy    1064 GTCAAAAGCTCTGCCCAGCCATGCACTCAACTCCATGAGCAGGAGGCTCCAGCCTCAAGAT 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1012 ATTTAAAACCTCTGCCCCAGCCACGCACTCACCTCTGTGAGCAGGGGTCCAGCCTCAAGAT 1071

Qy    1124 CCTTTCCAAAGGAAAGCGGGGGTGGACACTCTTCCGTCTCCGCCGAGTCAGAATCCTCCAG 1183
           ||||||||||||||||||||| ||||||||||| |||||||||||||||||||| |||
Db    1072 CCTCTCCAAAGGAAAGCGAGGTGGACATTCATCGTTTCCACTGAGTCTGAGTCTTCAAG 1131

Qy    1184 TTTCACTCCAGCTAACCCTATGCAAAGACTTATATATATATATATATGATGATATGATAAAA 1243
           ||||||||||||||||||||||||||||||||||||||||||||||||------TTTATACGATAAATA 1183
Db    1132 TTTCACTCCAGCTAACACAGATGTAAAAGACTTTTT------TTTATACGATAAATA 1183

Qy    1244 GAACTTTTTTATGTTACACATTTTCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT 1303
           ||| |||||||||||||||  ||||||||||||||||||||||| ||||| |||||||
Db    1184 ACTTTTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAAATTGTACAGTTTT 1243

Qy    1304 TTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTTAGTTTTTGTGAGGTTTGACT 1363
           |  ||||||||||||||||||||----CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATT 1297
Db    1244 TATTGCTTGTTGGATTTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATT 1297

Qy    1364 TAATTATATAAATAATATTGTTTTTTGTTTGTTCATGTGAATGAGCGTCTAGGAGCAGGACCT 1423
           |||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db    1298 GACTTATTTATA---TAAATTTTTTTTGTTCATATTGATGTGTGTCTAGGGAGGACCT 1353

Qy    1424 GTGGCCAAGTTCTTAGTAGCTGTTTATCGTGTGTAGGACTGTAGAAGTGTAGGAGGAGA 1483
           |||||||||||||||||||||||||||||||||||||||||||||||||||||------
Db    1354 GTGGCCAAGTTCTTAGTTCGTGCTGTATGTCTCGTGGTAGGACTGTGAGA------AAGGG 1405

Qy    1484 AACTGAACATTCCAGAATGTGTGAAATTTTGTGTAAGGCTGTCCA 1543
           |||||||||||||||||||||||||||||||||||||||||||||
Db    1406 AACTGAACATTCCAGAGCCGTGTAGTGAATCACGTAGTGAATCAGACTAAAGCTAGAAAATGATCCCCAGCTGTT 1465

Qy    1544 GCTGCATA 1551
           ||||||||
Db    1466 TATGCATA 1473
```

```
RESULT 12
AB268886
ID   AB268886   standard; cDNA; 1679 BP.
XX
AC   AB268886;
XX
DT   28-MAY-2003   (first entry)
XX
DE   Nucleotide sequence of human chemokine receptor CXCR4.
XX
KW   Human; chemokine receptor; CXCR4; viral infection; surface protein;
KW   respiratory virus infection; respiratory syncytial virus infection;
KW   RSV infection; bronchiolitis; bronchitis; pneumonia; asthma; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key                Location/Qualifiers
```

```
FT   CDS             89..1147
FT                   /*tag= a
FT                   /product= "CXCR4"
XX
PN   WO2003014153-A2.
XX
PD   20-FEB-2003.
XX
PF   12-AUG-2002; 2002WO-CA001248.
XX
PR   10-AUG-2001; 2001US-0311088P.
XX
PA   (TOPI-) TOPIGEN PHARM INC.
XX
PI   Renzi P, Zemzoumi K;
XX
DR   WPI; 2003-256541/25.
DR   P-PSDB; ABP97733.
XX
PT   Modulating viral infection of a cell, for treating or preventing
PT   respiratory virus infections, bronchitis, pneumonia or asthma, by
PT   modulating a binding interaction between a cell chemokine-receptor and a
PT   surface protein of the virus.
XX
PS   Disclosure; Page 117-119; 120pp; English.
XX
CC   The present sequence encodes human chemokine receptor CXCR4. The
CC   specification describes a method for modulating viral infection of a
CC   cell. the method comprises modulating a binding interaction between a
CC   cell chemokine-receptor and a surface protein of the virus. The proviso
CC   is that the cell chemokine-receptor is not CX3CR1 and that the virus is
CC   not HIV. The method is useful for treating or preventing respiratory
CC   virus infection in vertebrates, more particularly respiratory syncytial
CC   virus (RSV) infections, and related diseases, e.g. bronchiolitis,
CC   bronchitis, pneumonia or asthma
XX
SQ   Sequence 1679 BP; 400 A; 407 C; 368 G; 504 T; 0 U; 0 Other;
```

Query Match   50.8%; Score 952.8; DB 7; Length 1679;
Best Local Similarity   80.2%; Pred. No. 4.1e-174;
Matches 1210;   Conservative 0;   Mismatches 252;   Indels 46;   Gaps 6;

```
RESULT 13
ABZ42642
ID   ABZ42642 standard; DNA; 1679 BP.
XX
AC   ABZ42642;
XX
DT   04-MAR-2003  (first entry)
XX
DE   Human CXC chemokine receptor 4 nucleotide SEQ ID NO:75.
XX
KW   G protein-coupled receptor, GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
KW   Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW   osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW   graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW   psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW   mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW   hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW   ulcer; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200261087-A2.
XX
PD   08-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US050107.
XX
PR   19-DEC-2000; 2000US-0257144P.
XX
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Burmer GC,  Roush CL,  Brown JP;
XX
DR   WPI; 2003-046718/04.
DR   P-PSDB; ABP81796.
XX
PT   New isolated antigenic peptides e.g., for G protein-coupled receptors
PT   (GPCR), useful for diagnosing and designing drugs for treating conditions
PT   in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT   autoimmune diseases.
XX
PS   Disclosure; Fig 1; 523pp; English.
XX
CC   The present invention describes antigenic peptides (I) comprising: (a)
CC   any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular G
CC   protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity or
CC   avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC   gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC   antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-related disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC   GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC   exemplification of the present invention
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

     Query Match             50.8%;  Score 952.8;  DB 7;  Length 1679;
     Best Local Similarity   80.2%;  Pred. No. 4.1e-174;
     Matches 1210;  Conservative    0;  Mismatches 252;  Indels  46;  Gaps    6;

Qy          44 GCAGGTGCAGGTAGCAGTGACCCTGCGGTTGGTGCTGGCGTTACCACCACGGCTG 103
                  |||||| | | |||||  |||  ||  | ||| | ||||   ||||  || |
Db          12 GCGGCAGCAGTGAGCAAAGTGACGCCGAGGGCCTGAGTGCTGCCAGTAGCCATCTG 71

Qy         104 TAGAGGCAGTGTTGCCATGGAACCGATCAGTGTGAGTARTATACACTTCTGATAACTACTC 163
                  |||| |            |  ||| ||||          |||||||||||||||||
Db          72 GAGAACCAGCGGTTACCATGGA----GGGGATCAGTATATACACTTCAGTAACTACAC 126

Qy         164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGACCCTGCTTCGGGGATGAAAA 223
                  |||||||||| ||| || ||||||||||||||||||||||||||||| ||||||||||
Db         127 CGAGGAAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy         224 CGTCCATTCAATAGGATCTTCCTGCCACCATCTACTTCATCATTCTTTGACTGGCAT 283
                  |||  || ||  |||||||||||||||||||||||||||||||||| ||  ||||||
Db         187 TGCTAATTCAATAAAATCTTCCTGCCACCATCTCCATCANTCTTCTTAACTGGCAT 246

Qy         284 AGTCGGCAATGATTGGTGTCATCCTGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC 343
                  | |||||||||| |||||||||||||||||||||||||||||||| || | ||||||||
Db         247 TGTGGGCAATGATTGGTGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGAC 306

Qy         344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTGTTGTCATCACACTCCCCTT 403
                  ||||||||| ||||||||||||||||||||||||||  || ||||||| ||||| ||||
Db         307 GACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCAGCGCTTCCCTT 366

Qy         404 CTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCGTCCA 463
                  |||||||||||| ||||| || |||||||||||||||||  || ||| || ||| ||||
Db         367 CTGGGCAGTTGATGCCGTGGCAAACGGTACTTTGGGAACTTCCATGCAAGGCAGTCCA 426

Qy         464 TATCATCTACACTCAACCTCACAGCAGGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
                  | ||||| ||||||||||||| |||||||  |||||||||||||||||||| ||||
Db         427 TGTCATCTACACAGTCAACCTCCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy         524 CCGGTACCTCGGCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
                  |||||||||| | ||||||||||||| |||||||||||||||||||||||| ||||||||
Db         487 CCGGTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy         584 AAAGGGCAGTCATGTGGGCGTCTGATCCAGCCCTCCTGACTATACCTGACTTCAT 643
                  || ||||| ||||| ||| |||||||||||||||||||| |||| |||||||||||
Db         547 AAAGGTGTCTATGTTGGCGTCTGACATCCCGCTGACATCCCGCTGACTATTCCCGACTTCAT 606

Qy         644 CTTTGCCGACGTCAGCAGGGGGACATCAGTCAGGGGGATGACAGCAGTACATCGTGACCG 703
                  |||| ||||  |                   |  |||| |||| |||| ||| ||| ||||
Db         607 CTTTGCCAACG--------TCAGTGAGGCAGATGACAGATATCTGTAGTACCG 651

Qy         704 CCTTTACCCGATAGCCTGTGATAGTGGATGGTGTGTTCAATTCCAGCATAATAAGGTGGGTCT 763
                  |||||||||| ||  ||||||||||| ||||||| | ||||||||||  |||| ||| |||
Db         652 CTTCTACCCCAATGACTTGTGGGTGTGTGTTCCAGTTTCAGCACACTCATGGTTGGCCTT 711

Qy         764 CATCCTGCCCGGCATCGTCATCCTCCTGTTACTGACATCATCTCAAGCTGTCACA 823
                  |||||||||||||| ||||||||| |||||| |||||| || ||||| ||| ||
Db         712 TATCCTGCCTGGTATTGTCATCCTGCTATTGCATCTGCTATTCATCTCCAAGCTGTCACA 771

Qy         824 CTCCAAGGCCACCAGAGGCGCAAGGCCCTCAAGACAGCAGTCATCCTCATCCTAGCTTT 883
                  |||||||||||||| |||| |||||||  |||||| |||| || |||||||||||| |||
Db         772 CTCCAAGGGCCACCACAGAAGCCAAGGCCCTCAAGACCAAGAGCCAGTCATCCTCATCCTGGCTTT 831

Qy         884 CTTTGCTGCTGCCCAGATTATGTGGGATCAGCATGTGCACAAGTGGATCTCCATCCTTTGGG 943
                  ||||| ||||| ||  |||| ||||||| ||| ||||||| |||| ||||||||||| ||||
Db         832 CTTCGCCTGTTGCTGCCTTCATCACTTCACAATTGGGACAGCCATGGGATCAGCATGTGCACCA 891

Qy         944 AGTCATCAAGCAAGGATGGACTTGCACAAGTGCACAGAGCATTGTGCACAGACTCCTCATCACAGA 1003
                  |||||| |||||||| | |  ||  ||| |||||  | || ||||||| ||| | ||||||||||||
Db         892 AATCATCAAGCAAGGGGTGTGAGTTTGAGAACACTGTGACACTGTGCACAGACTCCTCATCATCACCGA 951

Qy        1004 GGCCTCGCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTGGGGGCCAA 1063
                  ||||| |||| | |||||||||| ||||||||||||||||| ||||| ||||| |||||
Db         952 GGCCCTAGCTTCTTCTTCCACTGTTGCTGAACCCCATCCTCTATGCCTTTCTTGGAGCCAA 1011
XX
```

DR   WPI; 2003-757037/71.
DR   P-PSDB; ADC98646.
XX
PT   Use of a CXC chemokine receptor 4 (CXCR4) inhibitor, for modulating
PT   cellular adhesion or inhibiting proliferation, movement, or morphological
PT   change in a small cell lung cancer (SCLC) cell population, or for
PT   treating a subject having SCLC.
XX
PS   Disclosure; SEQ ID NO 1; 78pp; English.
XX
CC   The invention relates to methods and compositions for inhibiting small
CC   cell lung cancer (SCLC) proliferation and metastasis through modulation
CC   of the activity or expression of CXC chemokine receptor 4 (CXCR4), and
CC   optionally, that of the tyrosine kinase receptor c-Kit. Administration of
CC   a CXCR4 inhibitor modulates cellular adhesion and inhibits cellular
CC   proliferation, motility, and morphological changes in an SCLC cell
CC   population. The invention is based on the finding that CXCR4 is
CC   ubiquitously expressed, and c-Kit is variably expressed, in SCLC cells.
CC   The invention also discloses methods of screening for CXCR4 inhibitors.
CC   determining whether a sample of lung cancer cells expresses CXCR4 and is
CC   therefore amenable to CXCR4 inhibitor therapy; determining whether a
CC   sample of lung cancer cells coexpress CXCR4 and c-Kit is is therefore
CC   amenable to treatment with a combination of a CXCR4 inhibitor and a c-Kit
CC   inhibitor; and a method of assessing whether CXCR4 inhibitor therapy or
CC   CXCR4 inhibitor/c-Kit inhibitor combination therapy should be continued,
CC   by determining CXCR4 (and optionally c-Kit) activity in two or more
CC   tumour samples taken from an SCLC patient during the course of treatment.
CC   The methods of the invention are useful for treating SCLC, for screening
CC   for agents for the treatment of SCLC, for determining whether a SCLC from
CC   individual patients is suitable for treatment with agents of the
CC   invention, and for determining whether such treatment is effective and
CC   should be continued. The present sequence represents cDNA encoding human
CC   CXCR4, which can be used in methods of the invention.
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   50.8%; Score 952.8; DB 9; Length 1679;
Best Local Similarity   80.2%; Pred. No. 4.1e-174;
Matches 1210; Conservative   0; Mismatches 252; Indels 46; Gaps   6;

```
Qy    44 GCAGGTGAGGTAGCAGTGACCCTCTGAGCGGTTGGTGTCCGGTAACCACACGGCTG 103
         ||||| ||||||||||| | |||| ||| ||||  || |||||| |||  |||| ||||
Db    12 GCGGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGTCTCAGTAGCCACCGCATCTG 71

Qy   104 TAGAGCGAGTGTTGCCATGGACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
         ||| || ||| ||||||||||      ||||||||| |||||||||||| |||||  ||||
Db    72 GAGAACCAGCGGTTACCATGGA-----GGGGATCCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAA 223
         |||| |||| ||| |||| ||||||||||||  ||||||||||  | ||  |||||||
Db   127 CGAGGAAATGGCCTCAGGGGACTATGACTCCATGAAGGACCCCGTTTCGTGAAGAAA 186

Qy   224 CGTCCATTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
         | | ||||||||| |||||||||||||||||||||||| ||||||||||  |||||||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCACCATCTACTTCCATCATCTTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGGCAATTGGTGATCCTGGTCATGGGTTACCAGAAGAGCTAAGGAGCATGAC 343
         | | ||||||||||||||||| |||||||||||||||||||||||    | || |||||||
Db   247 TGTGGGCAATGGCAATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTGCTGCTCACTCCCTT 403
         |||||||||| |||||||||||||||||||  ||| || ||  |||||||||
Db   307 GGACAAGTACAGGCTGCACCTGTCAGTGGCCGAGCTCCTGTTTGTCATCACACTCCCTT 366

Qy   404 CTTGGGCAGTTGATGCGATCGCCATGGCTGACTGGCTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
         ||||||||||||||||||  |  |     |   ||||||| ||||||||   | || |||| |||||||
Db   367 CTTGGGCAGTTGATGCGTGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTCTACAGACGCGTTCTCATCCTGGCCTTCATCAGCCTGA 523
         | |||||||||| ||||||||||||||| ||| ||||||||||||||||||||| ||||
Db   427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTTGCCCATTGTCCAGCGCCACCAACAGTCAACAGTCAAAGGCCAAGGAAACTGTCGGCTGGCTGA 583
```

Qy 1064 GTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATGAGCAGAGGCTTCCAAGAT 1123
         |||| ||| | |||||||||| |||| || | ||||||||||| | |||||||
Db 1012 ATTTAAAACCTTCTGCCCAGCACACGCACTCACCTCTGTGAGCAGAGGGTCCAGCCTTCCAAGAT 1071

Qy 1124 CCTTTCCAAAGGAAAGGCGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAG 1183
         || | |||||||||||| | |||||| | |||| |||||| |||| | |||||
Db 1072 CCCTCTCCAAAGGAAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTTCAGTCTTCAAG 1131

Qy 1184 TTTTCACTCCAGCTAACCCTTATGCAAAGACTTATATAATATATATATATGATAAA 1243
         |||||||||||||||| | ||| |||||| | | | |||||| |||||
Db 1132 TTTTCACTCCAGCTAACACAGATGTAAAAGACTTTT------TTTATACGATAAATA 1183

Qy 1244 GAACTTTTTATGTTACACATTTCCAGATATAAGAGACTCTTGTACAGTTTT 1303
         | ||||||| |||||||||||| |||||||||| |||| | | ||||||||
Db 1184 ACTTTTTTTAAGTTACACATTTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTT 1243

Qy 1304 TTTTTTTTTTTAATGACTGTTGGGAGTTTATGTTCCTTCAGTTTTGTGAGGTTTGACT 1363
         ||| ||| | ||||| ||||| | ||||| |||||||||||| ||| |||
Db 1244 TATTGCTTGTGTGGATTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATT 1297

Qy 1364 TAATTTATATAAATATATTGTTTTTGTTGTTTCATGTGAATGAGCCTCTAGGCAGGACCT 1423
         | |||||| | ||| | |||||||||||||| | | |||||| | ||| |||||||||
Db 1298 GACTTATTTATA---TAAATTTTTTTGTTTCATTATTGATGTGTGTCTCAGGCAGGACCT 1353

Qy 1424 GTGGCCAAGTTCTTAGTAGCTGTTATCGTGTGTAGGACTGTGTAGGAACTGTAGAGGAGA 1483
         |||||||||||||||| ||| || | | | |||||||||| || ||| |||| |
Db 1354 GTGGCCAAGTTCTTAGTTGCGTATGCTCTGTGGTAGGACTGTAGAA-------AAGGG 1405

Qy 1484 AACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCCTGGATCCTCAGCTGTT 1543
         |||||||||||||||| | ||| || | |||||||||| || |||||||||||||||
Db 1406 AACTGAACATTCCAGAGCGGTGTAGTGAATCACGTAGAGAATAAGCTAGAAAATGATCCCCAGCTGTT 1465

Qy 1544 GCTGCATA 1551
         | ||||||
Db 1466 TATGCATA 1473

RESULT 14
ADC98645
ID   ADC98645 standard; cDNA; 1679 BP.
XX
AC   ADC98645;
XX
DT   01-JAN-2004 (first entry)
XX
DE   Human CXC chemokine receptor 4 (CXCR4)-encoding cDNA.
XX
KW   Human; CXC chemokine receptor 4; CXCR4; CXCR4 inhibitor;
KW   small cell lung cancer; SCLC; cellular adhesion; cellular proliferation;
KW   metastasis; motility; morphological change; drug screening; monitoring;
KW   cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   89..1147
FT   /*tag= a
FT   /product= "Human CXCR4"
XX
PN   WO2003079020-A2.
XX
PD   25-SEP-2003.
XX
PF   20-MAR-2003; 2003WO-EP002916.
XX
PR   20-MAR-2002; 2002US-0366370P.
XX
PA   (DAND) DANA FARBER CANCER INST INC.
PA   (NOVS) NOVARTIS AG.
XX
PI   Salgia R;
XX

ABZ35348
ID   ABZ35348 standard; cDNA; 1711 BP.
XX
AC   ABZ35348;
XX
DT   05-FEB-2003 (first entry)
XX
DE   Human gene expression profile polynucleotide SEQ ID NO 459.
XX
KW   Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW   bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW   tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW   gene expression; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200274979-A2.
XX
PD   26-SEP-2002.
XX
PF   20-MAR-2002; 2002WO-US008456.
XX
PR   20-MAR-2001; 2001US-0276947P.
XX
PA   (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI   Wan J, Wang Y;
XX
DR   WPI; 2002-740862/80.
XX
PT   New gene expression profile generated from primary, endothelial,
PT   epithelial, and muscle cell types, useful for identifying disease
PT   pathologies involving alterations of gene expression, e.g. cancer.
XX
PS   Disclosure; Page 609; 850pp; English.
XX
CC   The invention relates to a gene expression profile comprising one or more
CC   genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
CC   is a coronary artery endothelium, umbilical artery or vein endothelium,
CC   aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC   endothelium, myometrium microvascular endothelium, keratinocyte
CC   epithelium, bronchial epithelium, mammary epithelium, prostate
CC   epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC   small airway epithelium, renal epithelium, umbilical artery smooth
CC   muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC   dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC   aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC   bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC   osteoblasts or prostate stromal cell. The gene expression profile is used
CC   for determining the level of RNA expression for a sample, determining the
CC   phenotype of a cell and distinguishing cell types. The gene or a protein
CC   expression profile is useful in identifying disease pathologies involving
CC   alterations of gene expression. The assessment of expression profiles may
CC   provide meaningful information with respect to tumour type and stage,
CC   treatment methods, and prognosis. The gene or protein expression profile
CC   may also be used for creating microarrays. The microarray is useful for
CC   genetic and physical mapping of genomes, DNA sequencing, genetic or
CC   medical diagnosis, genotyping of organisms, confirming cell or tissue
CC   identification and in identifying promising antibiotics, antiviral or
CC   antifungal agents
XX
SQ   Sequence 1711 BP; 423 A; 407 C; 373 G; 507 T; 0 U; 1 Other;

Query Match   50.8%;   Score 952.8;   DB 6;   Length 1711;
Best Local Similarity   80.2%;   Pred. No. 4.2e-174;
Matches 1210;   Conservative   0;   Mismatches 252;   Indels   46;   Gaps   6;

```
Qy     44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCCCTTTGGTGCTCCGGTAACCACCACGGCTG 103
          |||||| ||||||||||||||||||| || | ||||  |  |||||| |||||||| ||||
Db     47 GCGGCAGCAGGTAGCAGAAGTGACGCCCGAGGGCCTCCAGTGCTCCCACCACCGCATCTG 106

Qy    104 TAGAGGCGAGTGGTTGCCATGGAACCGATCAGTGTGGAGTATATACACTTCTGATAACTACTC 163
          ||||||||||||| ||| |||||||| ||| |||||||||| ||||||||| |||||| || ||
```

```
Db    487 CGGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGGAGGAAGCTGTTGGCTGA 546
Qy    584 AAAGGCAGTCTATGTGGGCGTCTTGATCCAGCCCTCCTGCACTATACCTGACTTCAT 643
Db    547 AAAGGTGGTTCTATGTTGGGCGTCTGGATCCCTGCGTCCCTGCTATTCCCGACTTCAT 606
Qy    644 CTTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCGTGACCG 703
Db    607 ..........TCAGTGAGCGACAGATGACAGATATCGTGACCG 651
Qy    704 CCTTTACCCGATGCCGTGGATGGTGGTGTTCAAATTCCAGCCATATAAATGGTGGGTCT 763
Db    652 CCTTCTACCCCAATGACTTGTGGGTGGTTGGTTGTGTTCAGCACACATCATGTTGGCCT 711
Qy    764 CATCCTGCCCGGCATCGTCATCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACA 823
Db    712 TATCCTGCCTGGTATTGTCATCCTGCTATTGTCATTATCATCTCCAAGCTGTCACA 771
Qy    824 CTCCAAGGGGCCACCAGAAGCGCAAGGCCCTCAAGACGACCAGTCATCATCCTCAGTTT 883
Db    772 CTCCAAGGGGCCACCAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCGGCTTT 831
Qy    884 CTTTGCCTGCTGGCTGCCATATTATGTGGGGATCACAGCATCAGCCTTCATCCTTTGGG 943
Db    832 CTTCGCCTGTTGGCTGCCTTACTACATTGGGATCAGCAGCATCAGCCTTCATCCTGA 891
Qy    944 AGTCATCAAGCAAGGATGTGACTTCGAGAGCAATGGACAAGTGGACTCCATCACAGA 1003
Db    892 AATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTGACAAGTGGATTTCCATCACCGA 951
Qy   1004 GGGCCTCCGCCTTCTTCCTCACTGTTGGCTGTGCCTGAACCCATCCTCTGGCCTTCTCGGGCCAA 1063
Db    952 GGCCCTAGCTTTCTTCCTCACTGTGTCTGAACCCATCCTCTATGCTTTCCTTGGAGCCAA 1011
Qy   1064 GTTCAAAAGCTCTGCCCAGCATGCACTCACTCAATGAGCAGAGGCTCCAGCCTCAAGAT 1123
Db   1012 ATTTAAAACCTCTGCCCAGCCACGCACTCACCTCTGTGAGCAGAGGGTCCAGCCTCAAGAT 1071
Qy   1124 CCTTCCAAAGGAAAGCCGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAG 1183
Db   1072 CCTCCCAAAGGAAAGCGAGGTGGACATTCATCTGTTCCACTGAGTCTTGAGTCTTCAAG 1131
Qy   1184 TTTTCACTCCAGCTAACCCTATGCAAAGACTTATATATATATATATGATAAA 1243
Db   1132 TTTTCACTCCAGCTAACACACAGATGTAAAGACTTTT-----TTTATACGATAAATA 1183
Qy   1244 GAACTTTTTTATGTTACACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT 1303
Db   1184 ACTTTTTTTAAGTTACACATTTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTT 1243
Qy   1304 TTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTGTGAGGTTTGACT 1363
Db   1244 TATTGCTTGTTGTTTTGGGATTTTTGT-----CTTGTGTTTTCTTTAGTTTTGTGAAGTTTAATT 1297
Qy   1364 TAATTATATAAATATTGTTTTTGTTGTTTTTCATGTGAATGAGCGGTCAGGCAGGACCT 1423
Db   1298 GACTTATTTATA---TAAATTTTTTTGTTTCATTATTGATGTGTGTCTCAGGCAGGACCT 1353
Qy   1424 GTGGGCCAAGTCTTAGTAGCTGTTTATCTGTGTGTAGGACTGTAGAGAGGAGA------ 1483
Db   1354 GTGGGCCAAGTCTTAGTTGCTGTATGTCTCGTGGGTAGGACTGTACAA------AAGGG 1405
Qy   1484 AACTGGAACATTCCAGAATGTGTGGGTAAATTGAAATAAAGCCGGTGATCCTCAGCTGTT 1543
Db   1406 AACTGAACATTCCAGAGCGTGTAGTGAATCACGTAAAGCTAGAAATGATCCCAGCTGTT 1465
Qy   1544 GCTGCATA 1551
Db   1466 TATGCATA 1473
```

RESULT 15

```
Qy 1244 GAACTTTTTATGTTACACATTTCCAGATAATAAGAGACTGACCAGTCTCTTGTACAGTTTT 1303
        |||||   ||||||| || ||||||||| ||||||||||||||||||  || ||||||||||
Db 1219 ACTTTTTTTTTAAGTTACACATTTTCAGATAATAAAAGACTGACCAATATTGTACAGTTTT 1278

Qy 1304 TTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTTGACT 1363
        |||||||| ||||  |||||||||||    ||| |||||||| |||||||||||
Db 1279 ----------CTTGTGTTGTTGGATTTTTGT------CTTGTGTTCTTTAGTTTTTGTGAAGTTTTAATT 1332

Qy 1364 TAATTTATATAAAATATTGTTTTTTTGTTGTTCATGTGTTCATGTGAATAGAGCGTCTAGGCAGGACCT 1423
        ||||||||||     |||||||||| |||||||||||||||||||
Db 1333 GACTTATTTATA---TAAATTTTTTTGTTTCATATTGATGTGTCTAGGCAGGACCT 1388

Qy 1424 GTGGCCAAGTCTTAGTAGCTGTTATCGTGTGTGTAGGACTGTGTAGGACTGTAGAGGAAGA 1483
        |||||||||| |||||| |||| | | |||| ||||||||||||||||||||
Db 1389 GTGGCCAAGTTCTTAGTTGCTGTATGCTCGTGTAGGACTGTAGAA------AAGGG 1440

Qy 1484 AACTGAACATTCCAGAATGTGTCGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTT 1543
        |||||||||||||||| ||||| | | ||||||| ||||||| | ||||||  |||||||
Db 1441 AACTGAACATTCCAGAGCCTGTGTAGTGAATCACGTAAAGCTAGAAATGATCCCCAGCTGTT 1500

Qy 1544 GCTGCATA 1551
        || |||||
Db 1501 TATGCATA 1508


Search completed: May 17, 2004, 06:37:02
Job time : 823.917 secs
```

```
Db  107 GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 161
        |||||||||||||||||||||       ||||||||||  ||||||||||||||
Qy  164 TGAAGAAGTGGGGTCTGAGACTATGACTTCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
        ||| |||| |||| ||| |||  |||| || ||||||||||||| || ||||||||||
Db  162 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCTGTTCCGTGAAGAAAA 221

Qy  224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCTTCATCATCTTCCTGACTGGCAT 283
        | || ||||||||  |||||||||||||||||||||||||||||||||||||   |||||||
Db  222 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCTTCATCATCTTCTTAACTGGCAT 281

Qy  284 AGTCCGCAATGGATTGGATCCTCGTCATGGGTTACCCAGAAGAGCTAGGGAGCATGAC 343
         ||| ||||||||||| | ||| ||||||||||||||||||||| |||| ||||||||
Db  282 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCCAGAAGAAACTGAGAAGCATGAC 341

Qy  344 GGACAAGTACCGGCTCACCTGCAGTGGCTGACCTCCCTTGTCATCACACTCCCTT 403
        |||||||||  |||| ||| | |||||||||||||||||||||||||| || ||||
Db  342 GGACAAGTACAGGCTGCACCGTGCAGTGGCTGACCTCCCTTGTCATCACGCTTCCCTT 401

Qy  404 CTGGGCAGTTGATGCCATGGCATGGTACTTTGGGAAATTTTGTGTTGTAAGGCTGTCCA 463
        |||||||||||||||| ||||    |||||||||||  ||  |||  ||||||| |||||
Db  402 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 461

Qy  464 TATCATCTACAACTGTGCAAACCTCTACAGCAGCGGTTCATCCTGGCCTTCATCAGCCTGGA 523
        | |||||||||  ||||||| || |||||||| || ||||||||||||||||||||| ||||
Db  462 TGTCATCTACACAGTCTCACAGCAGTGCCTCATCCTGGCCTTCATCAGTCTGGA 521

Qy  524 CCGGTACCTGCCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGTTGGCTGA 583
        ||| ||||||||||| ||||||||||||||||||||  |||||||||| |||||||||||
Db  522 CCGCTACCTGCCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 581

Qy  584 AAAGGCCAGTCTATGTGGGCGTCTGGATCCCCTCCTCCTCCTGACTATACCTGACTTCAT 643
        ||||| | ||| |||| ||| ||||||||||| ||| || ||  ||| |||||||||||||
Db  582 AAAGGTGTCTATGTTGGCTCTGGATCCCCTGCTATTCCCGACTTCAT 641

Qy  644 CTTTGCGACGTCAGCGCAGGGGACATCAGGGGGATGACAGAGTACATCTGTGACCG 703
        |||||||||    |||||||||||||    ||||| || ||| | |||||||||||||
Db  642 ------------TCAGTGAGGCAGATGACAGATATATCTGTGACCG 686

Qy  704 CCTTTACCCCGATAGCCTGTGATGGTGGTGTTCAATTCCAGATATAATAATGGTGGGTCT 763
        || || || |||| ||||| | ||| || | |||||||  |||| |||||||||||||||
Db  687 CTTCTACCCCAATGACTTGTGTGGGTGGTGTGTTCAGTTCGCAGCACATCATGGGTTGGCCT 746

Qy  764 CATCCTGCCCGGCATGTCATCCTCTCCTGTTACTGTGCATCATCATCCTAAGCTGTCACA 823
        || || ||||| ||| || ||||| || ||||| ||||| |||||| ||||||||| |||
Db  747 TATCCTGTCCTGTATTGTCATCCTGTCTGTTATTGCATTATCATCTCCAAGCTTGTCACA 806

Qy  824 CTCCAAGGGCCACCAGAGCGCAAGGCCCAAGGACGCAGTCATCCTCATCCTCTAGCTTT 883
        |||||||| |||||| |||  |||||||||||||| |||||||||||||| |||| ||||
Db  807 CTCCAAGGGCCACCAGAGAGCAAGGCCCAAGGCCCACCACAGTCATCCTCATCCTCTGGCTTT 866

Qy  884 CTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGCATTGTGCACAAGGATGACTTTTGGG 943
        ||||||| ||||||||| ||||||||| || ||| ||| || ||||| || || |||||||
Db  867 CTTCGCCTGTTGGCTGCCTTACTACACGTGTGGGATCAGCATGGTTGAGTTTGAGCATTGGG 926

Qy  944 AGTCATCAAGAGACAAGGATGTGACTTGCACTTCGAGAGACTTGTGCACAAGTGATCTCCATCACCAGA 1003
        ||||||||| |||||||| || || ||| |||| |||| ||||||||| |||||||||||||||||
Db  927 AATCATCAAGGATCAAGGTGTGAGTTTGAGAGTTGAGAACACTGTGCACAAGTGATTTCCATCACCGA 986

Qy 1004 GGCCCCTGCCCTTCTTCCACTGTTGCCTGAACCCCATCCTCCTATGCCTTCCTCCGGGGCCAA 1063
        ||| | || |||||||||||| || |||||||||||| |||| |||| || ||||| |||||
Db  987 GGCCTAGCTTTTCTTCCACTGTGTCTCACTGTTGCCTGAACCCCATCCTCTATGCTTTCCTTGGAGCCAA 1046

Qy 1064 GTTCAAAAGCTCTGCCCAGCACTCAACTCAACTGATGCACTCAATGCTCCATGCAGAGAGGCTCAAGAT 1123
        || ||||||| || || ||||||||| |||||| || || |||| ||||| ||| |||||| |||||||
Db 1047 ATTTAAAACCTCTGCCCAGCACGCACTCACCTGTGTCTGAGCACTCACGACCTCTGTGGAGCCTCAAGAT 1106

Qy 1124 CCTTTCCAAAGGAAAGCGGGGTGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAG 1183
        ||||| || ||||||||| || |||| ||||| |||||||||||||| | || |||| |
Db 1107 CCTCTCCAAAGGAAAGCGAGGAGGTGGACACTGTTTCCACTGAGTCTGTGAGTCTTCAAG 1166

Qy 1184 TTTTCACTCCAGCTAACCCTTATGCAAAGACTTATATAATAATATATATATATGATAAA 1243
        ||||||||||||| |||| |||||| || || |  ||||||  |||||||||||||||
Db 1167 TTTTCACTCCAGCTAACACAGATGTAAAAGACTTTTT------TTTATACGATAAATA 1218
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:52:09 ; Search time 150.281 Seconds
(without alignments)
6931.281 Million cell updates/sec

Title: US-09-367-052-1
Perfect score: 1877
Sequence: 1 ccatcctaatacgactcact........aaaaaaaaaaaaaaaaaa 1877
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	974.4	51.9	1737	1	US-08-202-056-4	Sequence 4, Appli
2	974.4	51.9	1737	1	US-08-076-093A-3	Sequence 3, Appli
3	974.4	51.9	1737	1	US-08-701-265-3	Sequence 3, Appli
4	974.4	51.9	1737	2	US-08-284-586-3	Sequence 3, Appli
5	974.4	51.9	1737	2	US-08-805-478-3	Sequence 3, Appli
6	974.4	51.9	1737	2	US-08-802-227A-3	Sequence 3, Appli
7	974.4	51.9	1737	2	US-08-801-238-3	Sequence 3, Appli
8	974.4	51.9	1737	2	US-08-801-228-3	Sequence 3, Appli
9	974.4	51.9	1737	3	US-09-104-296-3	Sequence 3, Appli
10	974.4	51.9	1737	5	PCT-US94-06380-2	Sequence 2, Appli
11	952.8	50.8	1679	4	US-09-517-605-14	Sequence 14, Appli
12	951.8	50.7	1664	4	US-09-582-224A-5	Sequence 5, Appli
13	951.8	50.7	1664	4	US-09-023-655-1213	Sequence 1213, Ap
14	859	45.8	1225	4	US-09-016-434-1235	Sequence 1235, Ap
15	834.6	44.5	1317	4	US-08-153-848-45	Sequence 45, Appli
16	834.6	44.5	1317	3	US-09-299-843A-45	Sequence 45, Appli
17	834.6	44.5	1317	4	US-09-088-337B-45	Sequence 45, Appli
18	834.6	44.5	1317	5	PCT-US93-11153-45	Sequence 45, Appli
19	198.4	10.6	1119	5	US-09-170-496D-65	Sequence 65, Appli
20	198.4	10.6	1679	1	US-08-202-056-6	Sequence 6, Appli
21	198.4	10.6	1679	1	US-08-076-093A-5	Sequence 5, Appli
22	198.4	10.6	1679	1	US-08-701-265-5	Sequence 5, Appli
23	198.4	10.6	1679	2	US-08-284-586-5	Sequence 5, Appli
24	198.4	10.6	1679	2	US-08-805-478-5	Sequence 5, Appli
25	198.4	10.6	1679	2	US-08-802-627A-5	Sequence 5, Appli
26	198.4	10.6	1679	2	US-08-801-238-5	Sequence 5, Appli
27	198.4	10.6	1679	2	US-08-801-228-5	Sequence 5, Appli
28	198.4	10.6	1679	3	US-09-104-296-5	Sequence 5, Appli
29	198.4	10.6	1679	5	PCT-US94-06380-3	Sequence 3, Appli
30	198.4	10.6	2818	3	US-08-982-493-7	Sequence 7, Appli
31	198.4	10.6	2818	3	US-08-628-655-1	Sequence 1, Appli
32	196.8	10.5	1119	3	US-09-170-496D-199	Sequence 199, App
33	195.8	10.4	2085	3	US-09-299-843A-65	Sequence 65, Appl
34	195.8	10.4	2085	4	US-09-088-337B-65	Sequence 65, Appl
35	193.8	10.3	1107	4	US-09-170-496D-19	Sequence 19, Appl
36	193.8	10.3	1293	4	US-09-016-434-1052	Sequence 105, Ap
37	193.8	10.3	1670	3	US-08-709-838-1	Sequence 1, Appli
38	193.8	10.3	1670	3	US-08-829-83-1	Sequence 1, Appli
39	193.8	10.3	1670	4	US-09-023-655-980	Sequence 980, App
40	193.8	10.3	1670	4	US-09-624-594-1	Sequence 1, Appli
41	192.2	10.2	1107	4	US-09-170-496D-173	Sequence 173, App
42	186.8	10.0	2751	1	US-08-153-848-23	Sequence 23, Appl
43	186.8	10.0	2751	3	US-09-299-843A-23	Sequence 23, Appl
44	186.8	10.0	2751	4	US-09-088-337B-23	Sequence 23, Appl
45	186.8	10.0	2751	5	PCT-US93-11153-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-202-056-4
; Sequence 4, Application US/08202056
; Patent No. 5440021
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Hebert, Caroline
; APPLICANT: Kim, Kyung Jin
; APPLICANT: Lee, James
; TITLE OF INVENTION: Antibodies to Human IL-8 Type B Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,056
; FILING DATE: 25-FEB-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-202-056-4

Query Match 51.9%; Score 974.4; DB 1; Length 1737;
Best Local Similarity 76.1%; Pred. No. 2.3e-195;
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;

Qy   50 GCAGTAGCAGTGACCCTCTGAGCGTTGGTGCTCCGGTAACCACCGGCTGTAGAGC 109
              ||||||| ||| |||| ||||| |||| || ||| ||
Db   20 GCGCGGCGCAAAGTGACCGCGAGGGCCTGAGCTGCTCCAGTGAGTTGCCACCGCATCTGGAGAAC 79

Qy   110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
              | ||||
Db   80 CAGGGGTTACCATGGA------GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134

Qy   170 AGTTGGGGTCTGGAGACTATGACCTCCAACAAGGAACCCGCTGCTCCGGATGAAAACGTCCA 229
              |||| | ||| ||| ||||||||| |||
Db   135 AATGGGCTCAGGGGACTATGACTCCATGGAAGGAACCCGTTTCCGTGAAGAAAATGCTAA 194

Qy   230 TTTCAATAGGATCTTCCTGCCCACCAACTCTTCATCATCTTCTGACTGGCATAGTCGG 289
              ||||||| ||||||||||||||||||||||||||
Db   195 TTTCAATAAAATCTTCCTGCCCCACCATCCATCATCTTCTAACTGGCATTGTGGG 254

Qy   290 CAATGGATTGGTGATCCTGGTCATGGGTTACCAGGAGCATGAGGAGCAA 349
              ||||||||||||||||||||||||||||||
Db   255 CAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAGAAGCAA 314

Qy   350 GTACCGGCGTGCACCTGTCAGTGGCTGACCTCCCTTGTCATCACACTCCCCTTCTGGGC 409
              ||||||||||||||||||||||||||||||||||||
Db   315 GTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACACGGTTCCCTTCTGGGC 374

Qy   410 AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
              |||||||||||| ||||||||||||||| ||||||||||
Db   375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGCAGTTCATGTCAT 434

Qy   470 CTACACTGTCAACCTCACAGCGCGTTCATCATCCTGGCCTTCATCAGCCTGACCGGTA 529
              ||| ||||||||||||||| |||||||||||||||
Db   435 CTACACAGTCAACCTCTACAGCAGTGTCCATCTGGCCTTCATGAGTCTGACCCGCTA 494

Qy   530 CCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGGCTGAAAAGGC 589
              |||||||||||||||||||||||||||||||||||||||||
Db   495 CCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGGGCTGAAAAGGT 554

Qy   590 AGTCTATGTGGGCGTCTGATCCCAGCCCTCCTGACTATACCTTGCATCTTCATCTTTGC 649
              |||| ||||||||||||||||||||||||||||||||||
Db   555 GGTCTATGTGGGCTGGATCCCCGTGCCCCTGCTATTCCCGACTATCCATCTTTGC 614

Qy   650 CGACCTCAGCCAGGGGACATCAGTCAGGGGGATGACAGTGACTACATCTGTGACCGCCTTTA 709
              | |||||||||| |||||||||||
Db   659 CAACG--------------TCAGTGAGGCAGATGACAGTGACATAT ATCTGTGACCGCCTTCTA 659

Qy   710 CCCCGATAGCCTGTGGATGGTGGTTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCT 769
              ||||||||||||||||||||||||||||| |||||||||||
Db   660 CCCCAATGACTTGTGGGTGTGTTGTTGTTTCAGTTTCAGTTATCAGCACATATGGTGGCCTTATCCT 719

Qy   770 GCCCGGCATCGTCATCCTCCTCCTGTACTGCATCATCATCTCTAAGCTGTCACACTCCAA 829
              ||||||||||||||||||||||||||||||||||
Db   720 GCCTGGTATTGTCATCCTGTCCTGTCTATTGCATTATCATCTCCAAGCTGTCACACTCCAA 779

Qy   830 GGGCCACCAGAAGCGCAAGGCCCTCAAGACAGCAGACATCATCCTCAGCTTTCTTTTGC 889
              |||||||||||||||||||||||||||||||||
Db   780 GGGGCACCACCAGAAGCGCAAGGCCCTCAAGACCATCATCCTCATCCTGGCTTTCTTCGC 839

Qy   890 CTGTGGCTGCCATATTATGTGGGGATCAGCATCGACTCCTTTTGGGAGTCAT 949
              | |||| |||| |||||||| ||| |||
Db   840 CTGTTGGCTGCCTTACTACATTGGGATCAGCATCGACTCCTTTATTCATCAT 899

Qy   950 CAAGCAAGGATGTGACTTCCAGAGACATTGTGCACAAGTGGATCTCCATCACAGAGGCCCT 1009
              ||||||||||||||||||||||| ||||||||||
Db   900 CAAGCAAGGGGTGAGTTTGAGAACACTGTGCACAAGTGGATTCCATCACCACCGAGGGCCCT 959

Qy 1010 CGGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTCGGGGGCCAAGTTCAA 1069
              |||||| |||||||||||| |||||| ||| |||||||||||||||||||||||
Db   960 AGCTTTCTTCCACTGTTGTCGAACCCCTTCTTCCTTGGAGCCAAATTTAA 1019

Qy 1070 AAGCTCTGCCCAGCATGCACTCAACTCAACTGGAGGCTCCAGCCTCCAGAGATCCTTTC 1129
              ||||||||||||||||||||||| |||||||||||||||||||||
Db 1020 AACCTCTGCCCAGCACGCACTCACCCTGGAGCAGAGGGTCCAGCCTCCAAGATCCTCTC 1079

Qy 1130 CAAAGAAAGCGGGGGTGGACACTCTTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTTCA 1189

Qy 1080 CAAAGGAAAGCGAGGTGGACATTCATCCTGTTTCCCACTGAGCTCTGTTTTCCACTGAGTCTGAGTCTGAGTCTGAAGTTTTCA 1139
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1190 CTCCAGCTAACCCTTATGCAAGGCTAAATAATAATATATAATATGATAAATAACTTT 1249

Qy 1190 CTCCAGCTAA-------CACAGATGTAAAAGACTTTTTTTTATACGATAAATAACTTTT 1191
              |||||||||| ||| |||||||||||||||||||||||||||||
Db 1140 CTCCAGCTAA-------CACAGATGTAAAAGACTTTTTTTTATACGATAAATAACTTTT 1191

Qy 1250 TTTTATGTTACACATTTTCAGATATAAGAGACTGACCAGTCTTGTCACAGTTTTTTTTTTT 1309
              ||||||||||||||||||||||||||||||||||||||||||
Db 1192 TTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTTTATTGC 1251

Qy 1310 TTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGTTTGACTTTAATTT 1369
              |||||||||||||| | ||| ||||||||||
Db 1252 TTGTTGGATTTTTGT-------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATTGACTTA 1305

Qy 1370 ATATAAATATTGTTTTTGTTTGTTCATGTGAATGAGCGTCTAGGCAGCAGGACCTGTGGCC 1429
              |||| | ||||||||||||||||||
Db 1306 TTTATA----TAAATTTTTTTGTTTCATATTGATGTGGTGTCTAGGCAGCAGCCTGTGGCC 1361

Qy 1430 AAGTTCCTAGTAGCTGCTGTTACTGTGTGTGGAGGACTGTAGAGACTGTAGAGGAAGAAACTGA 1489
              |||| ||| || |||||| ||||||||||||||||| |||||||||||
Db 1362 AAGTTCCTTAGTTGCTGTATGTCTCGTGGTAGGACTGTGAGAA---------AAGGGACTGA 1413

Qy 1490 ACATTCCAGAATGTGTGGTAAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCA 1549
              |||||||||||||||||||||||||||||||||||||||
Db 1414 ACATTCCAGACGGTGTAGTGAATCACGTAAAGCTAGAAATGATCCCCAGCGTGTTTATGCA 1473

Qy 1550 TAATCTCTTCATTCCGAGGAGCACCCACCCCCACCCCCACCCCCATTCTTAAAT 1609
              |||||||||||||
Db 1474 TAGATAATCTCTCCA-- 1488

Qy 1610 TGTTTGGTTATGCTGTGTGATGGTTTGTTTGTGTTTTTGTTGTGTGTTTTT 1669
              | ||||||||||||| |||||| |||||||
Db 1489 --------------------------TTCCCGTGGAACGTTTTTCCTGTCTTCTTAAGACGTGAT 1525

Qy 1670 TTTTCTGTAAAGATGGCACTTAAAACCTAAAAGCCTAAAAATGGTGGTAGAAATGCTGGGGT 1729
              |||| || ||| |||||||||||| |||| ||||||||||||
Db 1526 TTGCTGTGAGAAGATGGGCACTTATAACCAAAGCCCAAAAGTCGT-ATAGAAATGCTGG--- 1581

Qy 1730 TTTTTTGTTGTTGTTTGTTTTTTTCAGTTTTCAAGAGTAGAGATTGACTTCAGTCCTACACAAAT 1789
              | |||||||||||||||||| |||||||||||
Db 1582 -----------------TTTTTCAGTTTTCAGGGAGTGGGT-GATTTCAGCACCTAC-AGT 1623

Qy 1790 GTACAGTCTTGTATTACATTGTTAATAAAGTCAATGATAAAACTTAAAAAAAAAAAAA 1849
              ||||||||||||||||||||||||||||| |||||||||| |||||||||||
Db 1624 GTACAGTCTTGTGATTAAGTTGTTAATAAAGTACATGTTAAACTTAAAAAAAAAAAAAAA 1683

Qy 1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
              ||||||||||||
Db 1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

RESULT 2
US-08-076-093A-3
; Sequence 3, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:
;   APPLICANT: Chuntharapai, Anan
;   APPLICANT: Lee, James
;   APPLICANT: Hebert, Caroline
;   APPLICANT: Jin Kim, K.
;   TITLE OF INVENTION: Antibodies to Human PF4A Receptors
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076,093A
FILING DATE: 11-Jun-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 nucleotides
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear

US-08-076-093A-3

Query Match   51.9%;   Score 974.4;   DB 1;   Length 1737;
Best Local Similarity   76.1%;   Pred. No. 2.3e-195;
Matches 1391;   Conservative   0;   Mismatches 301;   Indels 136;   Gaps   10;

```
Qy      50  GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGTCCGGTAACCACCACGGCTGTAGAGC 109
                |||||||||||||||||||||||| |||||||||||||||||||||||||||
Db      20  GCGCGGCGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTGCCACCGCATCTGGAGAAC 79

Qy     110  GAGTGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAAACTACTCTGAAGA 169
                |||||                |||||||||||||||||||||
Db      80  CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134

Qy     170  AGTGGGGTCTGAGACTATGACTCCAACAACAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
                |||||||||||||||||||||||||||||||||||
Db     135  AATGGGCTCAGGGGACTATGACTCCATGAGCTCCAACAAGGAACCCGTTTCCGTGAAGAAAATGCTAA 194

Qy     230  TTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTCTGACTGGCATAGTCGG 289
                ||||||||||||||||||||||||||||||||||||||||||||
Db     195  TTTCAATAAAATCTTCCGCCACCATCTACTTCCATCATCTTCTTAACTGCCATTGTGGG 254

Qy     290  CAATGGATTGGTGATCCTCGGTCATGGGTTACCAGAAGAGAAGCTAAGGAGCATGACGGACAA 349
                |||||||||||||||||||||||||||||||||||||||
Db     255  CAATGGATTGGTCATCCTCGGTCATGGGTTACCAGAAGAGAAACTGAGAGACGAGCATGACGGACAA 314

Qy     350  GTACCGGCTGCACCTGTCAGTGGCTGACCACCTCCCTTCTTGTCATCAACACTCCCCTTCTGGGC 409
                ||||||||||||||||||||||||||||||||||||||
Db     315  GTACAGGCCTGCACCTGTCAGTGGCCGACCTCCCTTCCTTTGTCATCACGGCTTCCCTTCTCGGGC 374

Qy     410  AGTTGATGCCATGGCCATGCTGTACTTGGGAAATTTTGTGTAAGGCTGTGTCCATATCAT 469
                ||||||||||||||||||||||||||||||||||||||||||
Db     375  AGTTGATGGCCGTGGCAAACTGGTACTTTGGGAACTTCCATACGCAGTCCATGTCAT 434

Qy     470  CTACACTGTCAACCTGTCACAGCAGCGTTCTCATCCTGGCCTTCATCAGTCTGGACCGGTA 529
                |||||||||||||||||||||||||||||||||||||||||||
Db     435  CTACACAGTCAACCTCACAGCAGGTGTCCTATCCTGGCCTTCATCGTCTGGACCGCTA 494

Qy     530  CCTCGGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAAACTGCTGGCTGAAAAGGC 589
                |||||||||||||||||||||||||||||||||||||
Db     495  CCTGGCCATCGTCCACGCCACCAACAGTCAGAGGGCCAAGGAAAGGT 554

Qy     590  AGTCTATGTGGGGTCTGGATCCCAGCCCTTCCTCCTGACTATACCTTGACTTCATCTTTGC 649
                ||||||||||||||||||||||||||||||||||||||||
Db     555  GGTCTATGTTGGCGTCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCATCTTTGC 614

Qy     650  CGACGTCAGCAGCAGGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCCTTTA 709
                ||||||||||||||||||||||||||||||
```

Db   615 |||||
            CAACG-

```
Qy     710  CCCGATAGCCTGTGGATGGTGGTGTTCAATCCAGCATATAAATGGTGGGTCATCCT 769
                ||||||||||||||||||||||||||||||||||||||||||||
Db     660  CCCAATGACTGTGGGGTGTGTTGTCCAGTTTCAGCATCATCATTGCCTTTATCCT 719

Qy     770  GCCCGGCATCGTCATCCTCGTCTATTGCATTGCATTATCATCTCCAAGCGTGCACTCTCCAA 829
                |||||||||||||||||||||||||||||||||||||||||||||||
Db     720  GCCTGGTATTGTCATCCTCGTCTATTGCATTATCATCTCCAAGCGTGCACTCTCCAA 779

Qy     830  GGGCCACCAGAAGGCAAGGCCTCAAGACGAGTCATCCTCATCCTCTTCTTTGC 889
                ||||||||||||||||||||||||||||||||||||||||
Db     780  GGGCCACCAGAAGGCCAAGGCCCTCAAGACCCTCAAGACCATCCTCGCTTTCTTCGC 839

Qy     890  CTGCTGGCTGCCATATTATGTGGGGATCCAGCATCGCATCGACTCCTTTTGGGAGTCAT 949
                ||||||||||||||||||||||||||||||||||||
Db     840  CTGTTGGCTGCCTTACTACACTGGGATCAGCATCGACTCCTTCATCACCTGGGAAAATCAT 899

Qy     950  CAAGCAAGGATGTGACTTGCGAGAGCCATGTGCACAAGTGGATCCATCACAGAGGCCCT 1009
                |||||||||||||||||||||||||||||||||
Db     900  CAAGCAGGGTGTGAGTTGTGAGACACTGTGCACAAGTGGATTTCATCACCGAGGCCCT 959

Qy    1010  CGCCTTCTTCCACTGTGCTGAACCCCATCCTCTATGCCTTCCTGCGGGCCAAGTTCAA 1069
                ||||||||||||||||||||||||||||||||||||||
Db     960  AGCTTTCTTCCACTGTGCTGAACCCCATCCTCTATGCTTCCTATGCTTCCTGGAGCCAAATTTAA 1019

Qy    1070  AAGCTCTGCCCAGCATGCACTCAACTCACTCCATGAGACAGAGGCTCAGCCTCAAGATCCTTTC 1129
                ||||||||||||||||||||||||||||||||||||||
Db    1020  AACCTCTGCCCAGCACGCACTCACCTCGTGAGCAGAGGCCAGCCTCAAGATCCTCTC 1079

Qy    1130  CAAAGGAAAGCGGGGTGGACACTCTTCCGTCCACGTGCTCAGAATCCTCCAGTTTTCA 1189
                ||||||||||||||||||||||||||||||||||||||
Db    1080  CAAAGGAAAGCGAGGTGGACATTGCACTCTGTTCCACTGACGTCTGAGTCTCCAAGTTTCA 1139

Qy    1190  CTTCCAGCTAACCCTATGCAAAGACTTATATAATATATATATATATATAGAAGAACTT 1249
                |||||    |||||||||||||||||
Db    1140  CTCCAGCTAA-------CACAGATGTAAAAGACTTTTTTATACGATAAATAACTTTT 1191

Qy    1250  TTTTATGTTACACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTTT 1309
                |||||||||||||||||||||||||||||||
Db    1192  TTTTAAGTTACACATTTTCAGATATAAAGACTGACCAATATTGTACAGTTTTTTATTGC 1251

Qy    1310  TTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTGACTAATTT 1369
                ||||||||||||||||||||||||||
Db    1252  TTGTTGGATTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATTGACTTA 1305

Qy    1370  ATATAAATATTGTTTTTGTTTGTTCATGTGAATGAGCGTCTAGGCAGACACCTGTGGCC 1429
                ||||||||||||||||||||||
Db    1306  TTTATA-----TAAATTTTTTTTGTTTCATATTGATGTGTCTCAGGCCAGGACCTGTGGGC 1361

Qy    1430  AAGTTCTTAGTAGCTGTTTATCTGTGTAGGACTGTGTAGGAACTGTAGAGAGGAAGAAACTGA 1489
                ||||||||||||||||||||||||||||
Db    1362  AAGTTCTTAGTTGCTGTATGTCTCGTGGTAGGACTGTGAGA--------AAGGGACTGA 1413

Qy    1490  ACATTCCAGAAGTGTCGTGGTAAATTGAATAAAGTAGCCGTGGTGTTGCTGCA 1549
                ||||||||||||||||||||||||||||||||||
Db    1414  ACATTCCAGAAGCGTGAGTAGTGAATCACGTAGAAATGATCCCCAGCTGTTTATGCA 1473

Qy    1550  TAATCTTCATTCCGAGGAGCACCCCCACCCCCACCCCCACCCCCATCTTAAAT 1609
                ||||||||||||||
Db    1474  TAGATAATCTCTCCA------------- 1488

Qy    1610  TGTTTGGTTATGCTGTGATGGTTTGTTTGGTTTTTTTTTTGTTGTTGTTTTT 1669
                |||||||||||||||||||||||||||||||
Db    1489  ------------------TTCCCGTGGAACGTTTTCCTGTTCTTAAGACGTGAT 1525

Qy    1670  TTTTCGTAAAGATGGCACTTAAAACCAAAGCCTTGAAATGGTGGTAGAAATGGTCGGGGT 1729
                ||||||||||||||||||||||||||
Db    1526  TTGCTGTAGAAGATGGCACTTATAACCAAAGCCCAAAGCTT-ATAGAAATGTCTGG--- 1581

Qy    1730  TTTTTTGTTTGTTTGTTTTTTCAGTTTTTCAGTTTTCAGCTTCAGTTGACTTGACCCGCCTTTA 1789
                |||||||||||||
```

```
Db   170  AGTGGGGCTCGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA  229
          ||||| ||||| ||||||||||||||||  |  || ||||||  ||||||| |||| ||
Qy   135  AATGGGCTCAGGGGACTATGACTCCATGACTCGTTCCGTGAAGAAAATGCTAA         194

Qy   230  TTTCAATAGGATCTTCCTGCCCACCATCATCTTCATCATCTTCTTGACTGGCATAGTCGG  289
          |||||||   ||||||||||||||||||||||||  ||||||| ||||||||||| ||||
Db   195  TTTCAATAAAATCTTCCTGCCCACCATCATCTTCCATCATCTTAACTGGCATTGTGGG    254

Qy   290  CAATGGATTGGTCGATCCTGGTCATGGTTACCAGAGAGAAGCTAAGGAGCATGACGGACAA  349
          ||||||||||||| |||||||||||||||||||||||||||  ||||| ||||||| ||||
Db   255  CAATGGATTGGTCATCCTGGTCATGGTTACCAGAGAGAAGCATGACGACGACGACACAA    314

Qy   350  GTACCGGCTGCACCTGCAGTGGCTGCGACCCTCCGTCTGTCATCACACTCCCCTTGGGC   409
          |  |  | |||  || |||||| ||  ||| ||  | |  || |||||  ||| |||||
Db   315  GTACAGCTGCACCTGCAGTGCGCCACCCGCACCTCCTTCTTTGTCATCACCGGTTCCTGGGC 374

Qy   410  AGTTGATGCCATGGCTGACTGGTACTTGGGGAAATTTTGTGTAAGGCTGTCCATATCAT   469
          |||||||| |||  | || |||||||| |||||   ||| || || |||| ||  ||||
Db   375  AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT  434

Qy   470  CTACACTGTCAACCTCTCACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTA 529
          |  ||  ||||| ||||||| |||| ||  |||||| || ||| |||||||||| |||  |
Db   435  CTACACAGTCAACCTCTACAGCAGTGCTCCTCATCCTGGCCTTCATCAGCCTGTGACCGCTA  494

Qy   530  CCTCGCCATTGTCCAGCCACCAAACGTCAAAGGCCAAGGAAACTGCTGGTCAAAAGGC   589
          |||  || |||| ||| |  ||   ||||  || ||  ||||  || || |||||||
Db   495  CCTGGCCATCGTCCAGCGCCACCAACAGTCAGGAGAAGCTGTGGCCAAGAAGGT     554

Qy   590  AGTCTATGTGGGGCTCGGATCCCAGCCTCCTCCGACTATACCTGACTTCATCTTTGC    649
          |||| |  |||| ||   || ||  ||  || ||  | |  || |||| ||||| ||
Db   555  GGTCTATGTTGGCGTGTTGGATCCCGCCTCCTGCTGACATTATTCCCGACTTCATCTTTGC 614

Qy   650  CGAGTCAGCCAGGGGGACATCAGTCAGGGGATGACAGCAGATATCGTGACCGCGCTTTA   709
          | | ||   |  ||||||||||  |  ||||| |  | | | || |||  ||||||| |
Db   615  CAACG-----------TCAGTGAGGCAGATGACAGATGACATATCGTGACCGCGCTTCTA  659

Qy   710  CCCGGATAGCCTGTGGATGGTGGTTGTTCAATTCCAGCATATAAATGGTGGTCTCATCCT  769
          |||  || | ||||| |||| |||| |||||| ||  ||| | || || ||| ||| ||
Db   660  CCCCAATGACCTGTGGGTGGTGGTTGTTCAGTTCCAGCATCATCATCTCGGCCTTATCCT   719

Qy   770  GCCCGGCATCGTCATCCTCCTCTGTTACTGCATCATCATCCTCTAAGCTGTCACACTCCAA  829
          ||||  |||  | || |||| |  ||| || | || || ||  ||  ||||||| ||  |
Db   720  GCCTGGCATTGTCATCCTCTGTGTCCTGCCATTGCATTACATCCTCCCAAGCTGTCACACTCCAA 779

Qy   830  GGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGCAGTCATCCTCTAGTCTTTCTTTGC  889
          || ||| | ||||||||||||||  ||| |  || || ||  || |||  ||| ||| ||
Db   780  GGGCCACCAGAAGCGCAAGGCCCTCAAGACCATCATCCTCTTCCTCCTGGCTTCTTCGC    839

Qy   890  CTGCTGGCTGCCATATATGTGGGGATCAGCATGCATCGACTCCTTTTGGGAGTCAT   949
          | |||||||| | | |||||  || |  || || ||| |||||||||||||||||
Db   840  CTGTTGGCTGCCTTACTACATTGGAATCAGCATCAGCATCCTCCTGGAAATCAT    899

Qy   950  CAAGCAAGGATGTGACTTCGAGAGCATTGTCCACAAGTGGAATCTCCATCACAGAGAGCCCCT 1009
          |||| ||||  |||||||||||||  || | |||| || |||  || || |||||||||||
Db   900  CAAGGAAGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGGATTTCCATCACACCGGAGCCCT  959

Qy  1010  CGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTCGGGGGGCCAAGTTCAA 1069
          |||||||| |||| |||| |||||||||||||||| || ||| | ||  ||  |||  ||||
Db   960  AGCTTTCTTCCACTGTTGTCTGAACCCCATCCTCTATGCTTTCCTTGGAGCCCAAATTTAA   1019

Qy  1070  AAGCTCTGCCCAGCACCATGCAGTCAACTCCATGAGCAGGCGCTCCAGCCTCAAGATCCTTTC 1129
          | |||| |||||| || | ||| || |||  | |||||  | || |||  || |||| |||
Db  1020  AACCCTGCCCAGCACCACTCACCTGTGTCAGCCTCACGAGCAGCAGGGTCCAGCTCCAAGATCCTCTC 1079

Qy  1130  CAAAGGAAGCGCGGGTGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTTCA 1189
          |||| |||| | |||||| ||| | |||| ||||  |  ||| |||| |||| ||| ||
Db  1080  CAAAGGAAGCGAGGGTGACATTGTTCCACTGGATTCGTGAGTCTGAGTCTTCAAGTTTTCA  1139

Qy  1190  CTCCAGCTAACCCTTATGCAAAGACTTATATATATATGATAAACTACTCTGAAGA 1249
          |||||| | |  |  ||  | ||||||| | ||| ||| |||| |  ||| ||
Db  1140  CTCCAGCTAA------CACAGATGTAAAAGACTTTTTTTTTATACGATAAAATAACTTTTT 1191
```

```
Db  1582  -----------TTTTCAGTTTTCAGGAGTGGTTGATTCAGCACCTAC-AGT 1623

Qy  1790  GTACAGTCTTGTGTATTACATTGTTAATAAAAGTCAATGATAAACTTAAAAAAAAAAAAAA 1849
          |||||||||||||||||| |||||||||||||| || |||||||||||||||||||||||
Db  1624  GTACAGTCTTGTGTATTAAGTTGTTAATAAAAGTACATGTTAAACTTAAAAAAAAAAAAAA 1683

Qy  1850  AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
          |||||||||||||||||||||||||||
Db  1684  AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

RESULT 3

```
Query Match             51.9%;  Score 974.4;  DB 1;  Length 1737;
Best Local Similarity   76.1%;  Pred. No. 2.3e-195;
Matches  1391;  Conservative  0;  Mismatches 301;  Indels 136;  Gaps  10;

Qy   50  GCAGTTAGCAGTGACCCTCGAGGCGTTTGGTGCTGCTGGCCGTTCGGTAACCACCGGCTGTAGAGC 109
         |||  | |  ||  | |   | |   ||  ||| | || ||  ||  |||||| |||||   ||
Db   20  GCGCCGGCGCAAAGTGACGCGCCGAGGGCCTGAGTGCTCCAGTAGCGCCACCCGCATCTGGAGAAC  79

Qy  110  GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
         | |  |||||||||| |   ||||||  ||| ||| ||||| ||||| |||  ||  |||
Db   80  CAGCCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTCAGATAACCGAGGA 134
```

```
APPLICATION NUMBER:  US/08/076,093A
  FILING DATE:  11-Jun-1993
  APPLICATION NUMBER:  07/810782
  FILING DATE:  19-DEC-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:  07/677211
  FILING DATE:  29-MAR-1991
ATTORNEY/AGENT INFORMATION:
  NAME:  Love, Richard B
  REGISTRATION NUMBER:  34,659
  REFERENCE/DOCKET NUMBER:  706P2
TELECOMMUNICATION INFORMATION:
  TELEPHONE:  415/225-5530
  TELEFAX:  415/952-9881
  TELEX:  910/371-7168
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH:  1737 nucleotides
    TYPE:  Nucleic Acid
    STRANDEDNESS:  Single
    TOPOLOGY:  Linear

US-08-284-586-3

Query Match             51.9%;
Best Local Similarity   76.1%;    Score 974.4;   DB 2;   Length 1737;
Matches 1391;   Conservative   0;  Mismatches 301;  Indels 136;  Gaps  10;
```

```
Qy     50 GCAGGTAGCAGTGACCCTCTGAGCGGTTTGCGTCGTCGGTAACCACCACCGGCTGTGAGAGC 109
          |||||| |||||||  ||||||  ||||||  |||||||||| |||||||||||||||||||
Db     20 GCGCGGCGCAAAGTGACGCCGAGGGCCTGAGCGTGCTTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy    110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATCACTTCTGATAACTACTCTCGAAGA 169
          |||||||||||||||||||||        |||||||||||||||||||||||||||||||||
Db     80 CAGCGGTTACCATGGA------GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134

Qy    170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
          |||||| ||||||||||||| ||||||||||| |||||||||||||||||||  ||||||
Db    135 AATGGGCTCAGGGACTATGACTCCACCAACAGTCAATGAAGGAACCCGTTTCCGTGAAGAAAATGCTAA 194

Qy    230 TTTCAATAGGATCTTCCTGCCACCATCTACTTCTTCTGACTGGCATAGTCGG 289
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    195 TTTCAATAAATCTTCCTGCCACCATCTACTTCTTCTTCTTAACTGGCATTGTGGG 254

Qy    290 CAATGGATTGGTGATCCTGGTCATGGTCATGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAA 349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 CAATGGATTGGTCATCCTGGTCATGGTTACCAGAAGAAGAAGCTGAGAAGCATGACGACGACAA 314

Qy    350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTTCTTTGTCATCACAACTCCCCTTCTGGGC 409
          ||||| |||||||||||||||| ||||||||||||||| |||||||||||||||| ||||||
Db    315 GTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCTGTCATCACGCTTCCCTTCTGGGC 374

Qy    410 AGTGATGCCATGGCCATTGTCTGACTGTACTTGGAAATTTTGTAAGGCTGTCCATATCAT 469
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    375 AGTTGATGCCCTGGCAAACTTGGTACTTTGGGAAACTTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy    470 CTACACTGTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTA 529
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 CTACACAGTCAACCTCTACAGCAGCGTGTCCTCATCCTCGGCCCTTCATCAGTCTGGACCGGCTA 494

Qy    530 CCTCGCCATTGTCCACGCCCACCACCAGTCAAAGGCCCAAGGAAACTGCTGCTGAAAAGGC 589
          ||||||||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||
Db    495 CCTGGCCATCGTCCACGCCACCAACAAGTCAAAGGCCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGT 554

Qy    590 AGTCTATGTGGGCGTCTGGATCCCAGCCCCTCCGACTATACCTGACTTCATCTTTGC 649
          ||||| |||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    555 GGTCTATGTTGGCGTCTGGATCCCTGGCCCCTCCGACTGACTTCCCGACTTCATCTTTGC 614

Qy    650 CGACTCAGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCGTGACCGGCCTTTA 709
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    615 CAACG-----------------TCAGTGAGGCAGATGACAGACAGATATATCGTGACCGGTCTTCATCCT 659

Qy    710 CCCCGATAGCCTGTGATGGTGTTTCAATTCAGCATATAAATGTGGGTCTCATCCT 769
          |||| |||||||||||| |||||| |||||||||||||||||| ||||||||||||
```

```
Qy   1250 TTTTATGTTACACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTTT 1309
          |||||  ||||||||  ||||||||| ||||||||||||||||||||||||||||||||
Db   1192 TTTTAAGTTACACATTTTCAGATATAAAGACTGACCAATATTGTACAGTTTTTATTGC 1251

Qy   1310 TTTTAATTGACTGTGGGAGTTATGTCCCTAGTTTTGTGAGGTTTGACTTTTTT 1369
          ||||           ||||| ||||||   ||||||||||||||
Db   1252 TTGTTGGATTTTGT----CTTGTGTTTCTTAGTTTTTGTGAGGTTTAATTGACTTA 1305

Qy   1370 ATATAAATATTGTTTTTTGTTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGGCC 1429
          ||||||      |  |||||||| ||||||||| ||||||||||||||||||||||||||
Db   1306 TTTATA----TAAATTTTTTTTGTTTTTCATATTGATGTGTGTCTAGGCAGGACCTGGCC 1361

Qy   1430 AAGTTCTTAGTAGCTGTTTATCTGTGTGTAGGACTGTAGGAGGAAGAAACTGA 1489
          ||||||||||||| || ||||||||||| |||||||||        |||||||||
Db   1362 AAGTTCTTAGTTGCTGTATGTCTCGTGTGTAGGACTGTAGAA------AAGGGAACTGA 1413

Qy   1490 ACATTCCAGAAATGTGTGTAAATTGAATAAAGTAGCCGTGATCCTCAGCGTGTTGCTGCA 1549
          |||||||||||| ||| ||||||||||||| ||| |||||||||||||||||||||||||
Db   1414 ACATTCCAGAGCGTGTAGTAGTAGAAAATGATCACGTAAAATGATCCCCAGCTGTTATGCA 1473

Qy   1550 TAATCTTCCTCATTCCGAGGAGCACCCCACCCCACCCCCATTCTTAAAT 1609
          ||      ||||||||||||||||||||||||| || |||||||||||
Db   1474 TAGATAATCTCTCCA----------------------------------- 1488

Qy   1610 TGTTTGGTTATGCTGCTGTGATGGTTGTTGGTTTTTTGTTTGTTGTTGTTTTT 1669
          ||  |||  |||   ||||| ||   ||||||||||||||||||||||||||||
Db   1489 ------------------TTTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGAT 1525

Qy   1670 TTTTCTGTAAAAAGATGGCACTTAAAACCAAAGCTGAAAATGTGGTAGAAAATGCTGGGGT 1729
          |||||||  ||||||||||||  |||||| |||||||||||||| |||| |||||||||||
Db   1526 TTTGCTGTAGAAGATGGCACTTATAACCAAGCCCCAAAGTGT-ATAGAAAATGCTGG---- 1581

Qy   1730 TTTTTTTGTTTGTTTTTGTTTTTTCAGTTTTTGTTTGATTGACTTCAGTCCCTACAAAT 1789
                    ||||||||| |||||||||||||| || |||||||||||||||
Db   1582 -------------TTTTTCAGTTTTGTTTGTTGATTTCAGCATGTTTCAGCACCTAC-AGT 1623

Qy   1790 GTACAGTCTGTTGTTATTACATTGTTAATAAAAGTCAATGATAAACTTAAAAAAAAAAAAA 1849
          |||||||||||||||||||| |||| |||| |||||||||    |||||| ||||||||||
Db   1624 GTACAGTCTGTTGTTATTAAGTTGTTAAAAGTACATGTTAAAACTTAAACTTAAAAAAAAAAAAA 1683

Qy   1850 AAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
          |||||||||||||||||||||||||||
Db   1684 AAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

```
RESULT 4
US-08-284-586-3
; Sequence 3, Application US/08284586
; Patent No. 5840856
; GENERAL INFORMATION:
;  APPLICANT:  Chuntharapai, Anan
;  APPLICANT:  Lee, James
;  APPLICANT:  Hebert, Caroline
;  APPLICANT:  Jin Kim, K.
;  TITLE OF INVENTION:  Antibodies to Human PF4A Receptors
;  NUMBER OF SEQUENCES:  6
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  Genentech, Inc.
;  STREET:  460 Point San Bruno Blvd
;  CITY:  South San Francisco
;  STATE:  California
;  COUNTRY:  USA
;  ZIP:  94080
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE:  WinPatin (Genentech)
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER:  US/08/284,586
;  FILING DATE:
;  CLASSIFICATION:  424
;  PRIOR APPLICATION DATA:
```

RESULT 5
US-08-805-478-3
; Sequence 3, Application US/08805478
; Patent No. 5874543
; GENERAL INFORMATION:
;   APPLICANT: Chuntharapai, Anan
;   APPLICANT: Lee, James
;   APPLICANT: Hebert, Caroline
;   APPLICANT: Jin Kim, K.
;   TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/805,478
;   FILING DATE: 25-Feb-1997
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/284586
;   FILING DATE: 10-AUG-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/076093
;   FILING DATE: 11-JUN-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/810782
;   FILING DATE: 19-DEC-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B.
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: P0706P2P1C1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/37-7168
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1737 base pairs
;   TYPE: Nucleic Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Linear
US-08-805-478-3

Query Match   51.9%; Score 974.4; DB 2; Length 1737;
Best Local Similarity 76.1%; Pred. No. 2.3e-195;
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;

Qy   50 GCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACCGGCGTGTAGAGC 109
               ||| |||| ||| | ||| || || ||| || | |||||||| ||| ||| ||||
Db   20 GCGCGGGCGCAAAGTGACGCGGCCTGAGTGCTCCAGTAGCCACCACCGGCCATCTGGAGAAC 79

Qy   110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAGA 169
                                           |||||||||||||||||||||||||||||
Db   80 CAGCGGGTTACCATGGA------GGGGATCCAGTATATACACTTCAGATAATACACCGAGGA 134

Qy   170 AGTGGGCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGATGAAAACGTCCA 229
                |||||| |||||||||||||||| |||||| ||||| ||||| ||||| ||| ||| |
Db   135 AATGGGCTCAGGGGACTCAGGGGACTATGACTCCATGAAGCAACTTGTTCCGTGAAGAAAATGCTAA 194

```
Qy  1310  TTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTTGACTTAATTT  1369
                               ||||||    |||||||||||  ||||||||||||||||
Db  1252  TTGTTGGATTTTGT-------CTTGTGTTTCTTTAGTTTTGTGAAGTTTAATTGACTTA    1305

Qy  1370  ATATAAATATTGTTTTTGTTGTTTCATGTGATTGATGTGTCTTCAGGCAGGACCTGTGGCC  1429
                                    ||||||  |||||||||||||||||||||
Db  1306  TTTATA----TAAATTTTTGTTTTGTTCATTGATGTGTCTTCAGGCAGGACCTGTGGCC    1361

Qy  1430  AAGTTCTTAGTAGCTGTTTATCTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGA     1489
                                |||||||||||||||||||||||      ||||||||||
Db  1362  AAGTTCTTAGTTGCTGTATGTCTGTGGTAGGACTGTAGAA------AAGGGAACTGA      1413

Qy  1490  ACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTGTTGCTGCA        1549
                ||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db  1414  ACATTCCAGAGCGTGTAGTGAATCACGTAAAGCTAGAAAATGATCCCCAGCTGTTTATGCA   1473

Qy  1550  TAATCTTCATTCCGAGGAGCACCCCACCCCACCCCCCATTCCTAAAT        1609
                |||||||||||||
Db  1474  TAGATAATCTCTCCA-------------------------------------------    1488

Qy  1610  TGTTTGGTTATGCTGTGTGATGGTTTGTTGGTTTTGTTGTTTGTTTTT        1669
                                       |||||||||||||||||||||
Db  1489  ------------------------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGAT   1525

Qy  1670  TTTTCTGTAAAAGATGGCACTTAAAACCAAAGCCAAAGGGT----ATAGAAATGCTGG--  1729
                |||||||||||||||||||||||||||||||||||  |||||||    ||||||||
Db  1526  TTTGCTGTAGAAGATGGACACTTATAACCAAAGCCAAAGTGGT-ATAGAAATGTCTGG--- 1581

Qy  1730  TTTTTTGTTTGTTTGTTTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAAT   1789
                                          |||||||||||||||||||||||||||
Db  1582  ------------------TTTTTCAGTTTTCAGGAGTGGGTGATTTCAGCACCTAC-AGT  1623

Qy  1790  GTACAGTCTTGTATTACATTGTTAATAAAAGTCAATGATAAACTTAAAAAAAAAAAAAAA   1849
                |||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db  1624  GTACAGTCTTGTATTTAAGTTGTTAATAAAGTACATGTTAAACTTAAACTTAAAAAAAAAA  1683

Qy  1850  AAAAAAAAAAAAAAAAAAAAAAAAAAAA  1877
                ||||||||||||||||||||||||
Db  1684  AAAAAAAAAAAAAAAAAAAAAAAAAAAA  1711
```

```
Qy  230  TTTCAATAGGATCTTCCTGCCCACCATCTACTCATCATCTTCTTGACTGGCAATAGTCGG   289
              ||||||||    |||||||||||||||||||||||||||||||||||
Db  195  TTTCAATAAAAATCTTCCTGCCCACCATCTACTCACTCTTCTTAACTGGCAATTGTGGG   254

Qy  290  CAATGGATTGGTGATCCTCGGTCGTCATGGGTTACCAGAAGAAGCTAAGGAGGACGATGAGGACGACAA   349
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  255  CAATGGATTGGTCATCCTCGGTCATGGGTTACCAGAGAGAGAAACTGAGAGAAGCATGATGACGACAA   314

Qy  350  GTACCGGCTGACACCTGTCAGTGGCTGACCTCCTTTGTCATCAGACACTCCCCTTCTCGGGC  409
              ||||||||||||||||||||||||||||||||||||||||||||||
Db  315  GTACAGGCTGACCTGTCAGTGGCCGACCTCCTTTGTCATCACGCTTCCCTTCTCGGGC   374

Qy  410  AGTTGATGCCATGGCTGACTGTACTTGGAAATTTTGTGTAAGGCTGTCCATATCAT   469
              |||||||||||||||||||||||||||||||||||||||
Db  375  AGTTGATGCCGGTGGCAAACTGGTACTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT   434

Qy  470  CTACACTGTCAACCTCTACAGCAGCCGTTCATCCTGGCCTTCATCAGCTGGACCGGTA  529
              ||||||||||||||||||||||||||||||||||||
Db  435  CTACACAGTCAACCTCTACAGCAGTCGCTCTCACAGTGCTGGACCGCTA    494

Qy  530  CCTCGGCCATTGTCCACGCCACCAAACTGCTTGGATCCGCAAGGAAAAGGC   589
              ||||||||||||||||||||||||||||||||||||||||||
Db  495  CCTGGGCCATCGTCCACGCCACCAAACAGTCAGAGGGCCAAGGAGGCCAAGGAAAAGGT   554

Qy  590  AGTCTATGTGGGGCGTCTGGATCCCAGCCCTCCTCGACTATACCTTGC   649
              ||||||||||||||||||||||||||||||||||||||||||||||
Db  555  GGTCTCATATGTTGGCGTCTGGATCCCTGCCCTCTGCGACTTCATCTTTGC   614

Qy  650  CGACGTCAGCCAGGGGGACATCAGTCAGGTACATCTGTGACCGCCTTTA   709
              ||||||||||||||||||||
Db  615  CAACG---------TCAGTGAGGCAGATGACATGACAGAGTATATCTGTGACCGCCTCTA   659

Qy  710  CCCCGATAGCCGTGGTGGTGTTGTCAATTCCAGCATATAATGTGGGTCTCATCCT  769
              |||||||||||||||||||||||||||||||||
Db  660  CCCCAATGACTTGTGGTGGGTGTGTTGTTGTTCAGCAGTTTGCCTTATCCT   719

Qy  770  GCCCGGCATCGTCATCCTCTCCGTTACTGACTGTCATCCTCTAAGCTGTCACACTCCAA   829
              ||||||||||||||||||||||||||||||||||
Db  720  GCCTGGTATTGTCATCCTGTCCTCGCTATTGCATTATCATCCTCCAAGCTGTCACACTCCAA   779

Qy  830  GGGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGTGCATCCTCATCCTTCGTTTCTTGC   889
              ||||||||||||||||||||||||||||||||||||||||||
Db  780  GGGGCCACCAGAAGCGCAAGGCCCTCAAGACGACCCACAGTGCATCCTCATCCTGCTTCTTCGC   839

Qy  890  CTGCTGGCTGCCATATATGTGGGGATCAGCAACGCATGCCTGACTCCTTTTGGGAGTCAT   949
              |||||||||||||||||||||||||||||||||||||||||
Db  840  CTGTTGGCTGCCTACTACATTGGGATCAGCAGCGCATGGCTCCTTCATCTCCTGGAAATCAT   899

Qy  950  CAAGCAAGGATGGATGTGACTCGAGAGCATTGTGCACAAGTGACTCCATCACAGAGGCCCT   1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  900  CAAGCAAGGGTGTGAGTTTGAGAACACTGCACAACTGGGATTTCCATCACCGAGGCCCT   959

Qy  1010  CGCCTTCTTCACTGTTGCCTGGACCTGTGAACCCCATCCTCTATGCCTTCCTCCGGGGCCAAGTTCAA   1069
              ||||||||||||||||||||||||||||||||||||||
Db  960  AGTTTCTTCACTGTTGCTGGACCTGTGAACCCCATCCTCTATGCCTTCCTCTGGGAGCCAAATTTAA   1019

Qy  1070  AAGCTCTGCCCAGCATGCACTCAACTCCATCGAGCAGAGGCTTCCAGCCTTCAAGATCCTTTC   1129
              ||||||||||||||||||||||||||||||||||||||||||
Db  1020  AACCTCTGCCCAGCACCAGCACCACCTGTGAGCAGAGGGTTCCAGCCTCCAAGATCCTCTC   1079

Qy  1130  CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCA   1189
              |||||||||||||||||||||||||||||||||||||||||||
Db  1080  CAAAGGAAAGCGAGGTGGACATTCATCTGTTCCACTGAGTCGAGTCTTCAAGTTTCA   1139

Qy  1190  CTTCCAGCTAACCCTTATGCTGCAAAGACTTATATATATATATATGATAAAGAACTT   1249
              ||        |
Db  1140  CTCCCAGCTAA------CACAGACTAAAAGTAAAAGACTTTTTTTATACCGATAAAAAATAACTTTT   1191

Qy  1250  TTTTATGTTACACATTTTCAGATATAAAGACTGACCAATATTGTACAGTTTTTATTGC   1309
              ||||||||||||||||||||||||||||||||||||||||||
Db  1192  TTTTAAGTTACACATTTTCAGATATAAAGACTGACCAATATTGTACAGTTTTTATTGC   1251
```

```
;      APPLICATION NUMBER: 07/810782
;      FILING DATE: 19-DEC-1991
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Love, Richard B.
;      REGISTRATION NUMBER: 34,659
;      REFERENCE/DOCKET NUMBER: P0706P2P1D2
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 415/225-5530
;      TELEFAX: 415/952-9881
;      TELEX: 910/371-7168
;      INFORMATION FOR SEQ ID NO: 3:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 1737 base pairs
;      TYPE: Nucleic Acid
;      STRANDEDNESS: Single
;      TOPOLOGY: Linear
US-08-802-627A-3

Query Match      51.9%;  Score 974.4;  DB 2;  Length 1737;
Best Local Similarity  76.1%;  Pred. No. 2.3e-195;
Matches 1391;  Conservative 0;  Mismatches 301;  Indels 136;  Gaps 10;
```

```
Qy    50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCAGGCTGTAGAGC 109
            |||||  ||||||||   ||| ||   ||||||||||  ||| |||||    ||| ||||||
Db    20 GCGCGGCGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCCGCATCTGGAGAAC 79

Qy   110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
            |||||  | ||||||||      ||||| |||||||||||||   ||||||||||   ||
Db    80 CAGCCGGTTACCATGGA-----GGGGATCAGTCAGTATATACACTTCAGATAACTACACCGGAGGA 134

Qy   170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
            |||||  || ||||||||||||||||  | |||||||| |||||| ||||||||  ||||
Db   135 AATGGGCTCACGGGACTATGACTCCATGAGGAACCCTGTTCCGTGAAGAAAATGCTAA 194

Qy   230 TTTCAATAGGATCTTCCTGCCACCATCTACTTCCTGCCACCATCATCTTCTTAACTGGCATTGTGGG 289
            ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   195 TTTCAATAAAATCTTCCTGCCACCATCTACTTCCTGCCACCATCATCTTCTTAACTGGCATTGTGGG 254

Qy   290 CAATGGATTGGTGATCCTGGTCATGGGTCATGGGTTACCAGAAGAGCTAAGGAGCATGACGGACAA 349
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   255 CAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAACTGAGAAGCATGAGACGACGACA 314

Qy   350 GTACCCGCTGCACCTGCAGTGGCTGACCTCCTTGTCATCACACACTCTCGGGC 409
            ||||||| |||||||||||||||  |||||||||||||||||||||||||||||
Db   315 GTACAGGCTGCACCTGCAGTGGCCGACCTCCTTGTCATCACGCTTCCCTTCTGGGC 374

Qy   410 AGTTGATGCCATGGCCGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
            ||||||||||||||||   ||| ||||||||||||| ||||  |||||||| |||||||
Db   375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTGGGAACTAGAAGCAGTCCATGTCAT 434

Qy   470 CTACACTGTCAACCTCACAGCAGCGTTCTCATCCTGGCCTTCATCCTGGACCGGTA 529
            ||||| |||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db   435 CTACACTCAACCTCACAGCAGCGTTCTCCTCATCCTGGCCTTCATCAGTCTGGACCGCTA 494

Qy   530 CCTCGCCATTGCCACGCCACCAACGCCAACAAGGCCAAGGAAAGCCAAGGAAAAGC 589
            ||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   495 CCTGGCCATCGTCCACGCCACCAACGCCACCAGTGAGCTGTTGGCTGAAAAGGT 554

Qy   590 AGTCTATGTGGGCGTCTGGATCCCAGCCCTTCTGCTCGACTATACCGGCCTTTGC 649
            ||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   555 GGTCTATGTGGGCGTCTGGATCCCCGCCCTCTGCTTGACTATTCCCGACTTCATCTTTGC 614

Qy   650 CGACGTCAGCCAGCCGGGGACATCAGGGGGATGACAGTACATCTGTGACCGCCTTTA 709
            ||||||||||||  ||||||||  |||||||  ||||  |||||||||||||||||||||
Db   615 CAACG----------TCAGTGAGGCAGCAGATGACAGATATCTGTGACCGCCTTCTA 659

Qy   710 CCCCGGATAGCCTGTGATGGTGGTGTTTCAATTCAGCAGTCATATATGTGGGTCTCATCCT 769
            ||||||||||||||||||||||||||| |||||||||||||||||||  || ||||||||
Db   660 CCCCAATGACTGTGGGTGGTGTGTTTCAGCAGTTTCAGTTTCATTATCATCTTCCAGCGTCACACTTCCAA 719

Qy   770 GCCCGGCAGTCGTCATCTGTTACTGCATCATCATCTCCTGTTACTGCACTCTTAAGCGTGTCAACTCCAA 829
            |||||||||||  |||||| ||||||||||||||| |||||||||||||
Db   720 GCCTGTGTATTGTCATCCTGTCCTGCTATTGCATCATTACATCTCCAAGCGTGTCAACTCCAA 779
```

```
Qy   830 GGGCCACCAGAAGCGCAAGGCCCTCAGACGACGACAGTCATCCTCATCCTAGCTTCTTTGC 889
            |||||||||||| ||||||||||||  |||  |||||||||||||||||| |||| || ||
Db   780 GGGCACCAGAAGGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTGGCTTTCTTCGC 839

Qy   890 CTGCTGGCTGCCATATATGTGGGAGGATCAGCATCGACTCCTCCTTCATCCTTTGGGAGTCAT 949
            ||| |||||||||  |  || ||||  |||||||||| |||||| |||| |||| ||||||
Db   840 CTGTTGGCTGCCTTACTACATTGGGATCATCAGCATCAGCATCCTCATCCTGAAATCAT 899

Qy   950 CAAGCAAGGATGTGACTTCGGAGAGCATTGTGCACAAGTGGATCCTCATCACAGAGCCCCT 1009
            |||||||||||||||  ||| |||||| ||  ||| |||||| ||||| |||||||||||
Db   900 CAAGCAAGGTGTGAGTTTGAGAACACTGTCGCACAAGTGGATTCCATCACCAGCGAGGCCCT 959

Qy  1010 CGCCTTCTTCCACTGTGCCTGAACCCATCCTCACTGAGCAGAGGCTCCAGCCTCAAGATCCTTTC 1069
            ||||||||||||||||||  ||||| |||| ||||||||||||| |||| |||||||||||| |
Db   960 AGCTTCTTCCACTGTGTCAACCTCCTCCATCCTGTGAGCAGGGTCCAGCGTCCAGATCTCCTC 1019

Qy  1070 AAGTCTGCCAGCATGCACTCAACTCCAATCCATGAGCAGAGGCTCAGCGCCTCAAGATCCTTTC 1129
            |||||| |||||||| ||||  |||| ||||||||||||| |||| |||||||| |||| ||||
Db  1020 AACCTCTGCCCAGCACGCACTACCCTCTGTGAGCAGGGTCCAGCCTCAGCCTCAAGATCCTCTC 1079

Qy  1130 CAAAGGAAAGCGGGGTGACACTCTTCCGTCCAGTAACCTCTATATAATATATATATTCCAGTTTCA 1189
            |||||||||||||||| |||||||  | || ||| |||  |||||||||||||||||||||||||
Db  1080 CAAAGGAAAGCGGAGGTGGACACTGCTGCGTGGGACTGAGTGTCACTGATTATTTTTATACAGTTTCA 1139

Qy  1190 CTCCAGTCAACCCTTATGCAAAGACTTATATAATAATATATATATGATAATATGATAAAGAACTT 1249
            |||||||  ||   | || ||| |||   |||||||||||||||  |||||||||||| |||| |
Db  1140 CTCCAGTCAA-------CACAGATGTAAAAGACTTTTTTTATACGATAAATAACTTTT 1191

Qy  1250 TTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAGTCTGTACAGTGTTTTTTTT 1309
            ||| |||||||||||||||||||||||  ||||||||||||||  |||||   || |||
Db  1192 TTTTAAGTTACACATTTTCAGATATAAAGAGACTGACCAAATTGTACAGTTTTTATTGC 1251

Qy  1310 TTTTTAATTGACTGTTGGGGAGTTTATGTTCCTTCAGTTTTTGTGAGGTTTGACTTAATTT 1369
            |||||||||||| ||| | |  || |||| ||| ||||||| |||| |||||  ||||| |
Db  1252 TTGTTGGGATTTTGT---------CTTGTGTTTCTTTAGTTTTGTTGGAAGTTTAATTGACTTA 1305

Qy  1370 ATATAAAATATTGCTTTTTGTTTTTGTTCATGTGAATGAGCGTCTAGGCAGACCTGTGGCC 1429
            ||||   |  || |||| |||||||  ||||| |||  |||  | |||| |||  ||||||
Db  1306 TTTATA-----TAAATTTTTTTGTTTCATATATGATGTGTGTCTAGGCAGGACCTGTGGCC 1361

Qy  1430 AAGTTCTTAGTAGCTGTTTATCTGTGTGGTAGGACTGTGTGAGAACTGTGTAGAGGAGAAACTGA 1489
            |||||||||||||||  |  ||||| |  |||||||| ||||| |          |||||
Db  1362 AAGTCTTAGTTTGCTGTATGTCGTGTGGTAGGACTGTAGAA------AGGGAACTGA 1413

Qy  1490 ACATTCCAGAATGTGTCGGTAATGATGAATAAAGCTACGCCGTGATCCTCAGCTGTTGCTGCA 1549
            |||||||||||| || |||||| || |||| ||| |||  | |||||||| |||| |||||||
Db  1414 ACATTCCAGAGCGTGTAGTCGTAGTGAATCACGTAGAAATGATCCCCAGCTGTTTATGCA 1473

Qy  1550 TAATCTCTTCATTCCGAGGAGCACCCCACCCCCACCCCCCATTCTTAAAT 1609
            ||| |  ||||| ||  ||  |||||| |||||||||||||||| ||||
Db  1474 TAGATAATCTCTCCA------------AGCACCCCACCCCCCATTCTAAAT 1488

Qy  1610 TGTTTGGTTATGCTGTGTGATGGTTGTTTGTTGTTTTTGTTGTTGTTTTT 1669
            |||||||||||| |||||||||| |||| ||
Db  1489 -------------------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGAT 1525

Qy  1670 TTTTCTGTAAAAGAATGGCACTTAAAACCAAAGCCTGAAAATGGTAGAAAATGCTGGGGT 1729
            || | | |  |||||  ||| ||||  || ||||||||| |  ||    |||  ||| ||
Db  1526 TTTGCTGTAGAAGATGGCACTTATAACCAAAGCCCAAGTGGT--ATAGAAATGCTGG---- 1581

Qy  1730 TTTTTTTGTTTGTTGTTTTTTCAGTTTTCAGTTGACTTCAGTCCCTACAAAT 1789
                     ||||||| || ||||||||   |||||| |||| |  ||  ||
Db  1582 ----------TTTTTTCAGTTTTTCAGGAGTGGGTTGATTTCAGCAACTAC-AGT 1623

Qy  1790 GTACAGTCTTGTATTACATTGTTAATAAAGTCAATGATAAAACTTAAAAAAAAAA 1849
            |||||||||||||||||||| ||||||||| ||| |||| |||| |||||||||||
Db  1624 GTACAGTCTTGTATTAAGTTGTTAATAAAGTACAATGTTAAACTTAAAAAAAAAAA 1683

Qy  1850 AAAAAAAAAAAAAAAAAAAAAAA 1877
            |||||||||||||||||||||
Db  1684 AAAAAAAAAAAAAAAAAAAAAAA 1711
```

RESULT 7
US-08-801-238-3
; Sequence 3, Application US/08801238
; Patent No. 5919896
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,238
; FILING DATE: 19-Feb-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear

US-08-801-238-3

Query Match   51.9%; Score 974.4; DB 2; Length 1737;
Best Local Similarity   76.1%; Pred. No. 2.3e-195;
Matches 1391; Conservative   0; Mismatches 301; Indels 136; Gaps 10;

```
Qy    50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC 109
            |||  |||  || |    | ||   |||  | |||||  ||  ||||||  |||  ||
Db    20 GCGCCGGCGCAAAGTGCACGCCGAGGGCCTCGAGTGCTCCAGTAGCCACCGCATCGGAGAAC 79

Qy   110 GAGTGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
           ||  |  ||||| ||||      || |||| ||| |||||||
Db    80 CAGCGGGTTACCATGA-----GGGGATCAGTAGTATACATTACACTTCAGATAATACACCGAGGA 134

Qy   170 AGTGGGGGTCTGGAGACTATGACTTCCAAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
           ||| ||| ||||||| |||||| | | ||   ||||    |||  || ||||| |||  |||
Db   135 AATGGGCTCAGGGGACTATGACTTCATGATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy   230 TTTCAATAGGATCTTCCTGCCCACCATCCTCACTTCATCATCATCCACCACCACCTCCA 289
           |||||||  | ||| ||||| ||||| |||||||| || ||||||| || ||| |  |
Db   195 TTTCAATAAAATCTTCCTGCCCACCACCATCCATCATCTTCCTTAACTGACGCATTGTGGG 254

Qy   290 CAATGGATTGGTGATCCTGTCATGGAGTACCAGAAGAGCTAAGGAGGACGATGACGACGACAA 349
           ||||||||||||||||||||| || ||      |  |  ||  |  |  | ||||   |||
Db                                                                   CAATGGATTGGTCATCCTGTCATGGGTTACCAGAAGAAGCATGAAGCATGACGACGACAA 314
```

```
Db   255 CAATGGATTGGTCATCCTGTCATGGGTTACCAGAAGAAGCATGAAGCATGACGACGACAA 314
Qy   350 GTACCGGCTGCACCTGTCAGTGGCTGACGACGCACCTCCTCCTTGTTGTCATCACACTTCTGGGC 409
Db   315 GTACAGGCCTCTGCAGTCTCAGTGCCGTCAGGCCCACCTCCTGCTCATCAGCACTTCATCTTTGC 374
Qy   410 AGTTGATGCCATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCGTCTCCATATCAT 469
Db   375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434
Qy   470 CTACACTGTCAACCTCTACAGCAGCGTCTTCATCCTGGCCTTCATCAGCCTGGACCGGTA 529
Db   435 CTACACAGTCAACCTCTACAGCAGCGTCTTCATCCTGGCCTTCATCAGTCTGGACCGCTA 494
Qy   530 CCTCGCCATTGTCCACGCCACCAACAGTCAAACAGTCAAGGCCAAGGAAAACGCTGCTGACTTCATCTTTGC 589
Db   495 CCTGGGCCATGTCCACGCCACCAACAGTCAGAGCCAAGGAAGCGTGTGGCTGAAAAGGT 554
Qy   590 AGTCATGTGGGCGTCTGGATCCCACCAGCCCTCCTCCGACTATACCTGACTTCATCTTTGC 649
Db   555 GGTCTATGTTGGCGTCTGGATCCCCTGGACCCTGACTATTCCCGACTTCATCTTTGC 614
Qy   650 CGACGTCAGCCAGGGGACATCAGTCAGGGGGACATCAGTACACATCGTGACCGCCTTTA 709
Db   615 CAACG-------TCAGTGAGGCAGATGACAGATATATCTGTACGCGCTCTA 659
Qy   710 CCCCGATAGCCTGTGATGGTGTGTGTTTCAATTCCAGCATATAATGGTGGTCTCATCCT 769
Db   660 CCCCAATGACTTGTGTGGTGGTTGTGTTTCCAGTTTCAGCACATCATGGTTGGCCTTATCCT 719
Qy   770 GCCGCGGATCGTCATCCTCTGTTACTGGCATCATCATCCTCTAAGCTGTCACACTCCAA 829
Db   720 GCCTGGTATTGTCATCCTCGTCTGCTATTGGCATTATCATCCTCCAAGCTGTCACACTCCA 779
Qy   830 GGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTCAGCTTTCTTTGC 889
Db   780 GGGCCACCAGAAGCGCAAGGCCCTCAAGACCCACCACCAGTCATCCTCATCCTCGGCTTTCTTCGC 839
Qy   890 CTGCTGCTGCCATATTATGTGGGATCAGACTGTGCACAAGTGATCTCCATCACAGAGGCCCTT 949
Db   840 CTGTTGGCTGCCTTACTACATTGGGATCAGCACTGTCCATCACTGATCTCCATCCTCCTGGAAATCAT 899
Qy   950 CAAGCAAGGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGAGGGCCCT 1009
Db   900 CAAGCAGGGTGGAGTTGAGTTGAACAACTGTGCACAAGTGGATTTCCATCACCGAGGGCCT 959
Qy  1010 CGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCATGGCCCTTCCTCCGGGGCCAAGTTCAA 1069
Db   960 AGCTTTCTTCCACTGTTGTCTGAACCCCATCCTCATGCTTTCCTTCCTTGGAGCCAAATTTAA 1019
Qy  1070 AAGCTCTGCCCGCCATGCCACTCAACTCCATGCTCATGAGCGAGGCTCCAGCCTCAAGATCCTTTC 1129
Db  1020 AACCCTGCCCAGCACCGCACTCACCCTGTGAGCAGAGGGTCCAGCCTCCAGCTCCAGCCTCAAGATCCTC 1079
Qy  1130 CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCCAGGAGTCAGAATCCTCCAGTTTCA 1189
Db  1080 CAAAGGAAAGGAGGTGAGGTGACATTCATCGTGTTCCACTGAGGTCGAGTCTTCAAGTTTTCA 1139
Qy  1190 CTCCAGGTAACCCTTATGCAAAGACTTATATATAATAATATATGAATAAAGACTT 1249
Db  1140 CTCCAGCTAA------CACAGATGTAAAAGACTTTTTTTTATAGATAAAATAACTTT 1191
Qy  1250 TTTTATGTTAACACATTTTCCAGATATAAGAGACTGACGAAAAACGTCCA 1309
Db  1192 TTTTAAGTTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTTATTGC 1251
Qy  1310 TTTTTAATTGACTGTTGGGAGTTATGTTCCTTCTAGTTTTTGTGAAGTTGAGGTTGACTTTAA 1369
Db  1252 TTGTTGGATTTTGT-------CTTGTGTTCTTTAGTTTTTGTGAATGAGGTCTAGGGCATTGTGGG 1305
Qy  1370 ATATAAAATATTGTTTTCATTTGTTGTTTCATGTGAATGAGTGTGTGTGTCTAGGCAGCAGGACCTGTCGCC 1429
Db  1306 TTTATA----TAAATTTTTTTTTGTTTCATATTGATGTGTGTGTCTAGGCAGCAGGACCTGTCGCC 1361
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-801-228-3

```
Query Match             51.9%;  Score  974.4;  DB 2;  Length 1737;
Best Local Similarity   76.1%;  Pred. No. 2.3e-195;
Matches 1391;  Conservative    0;  Mismatches  301;  Indels  136;  Gaps  10;

Qy    50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACACGGCGTGTAGAGC 109
            ||||  ||   ||||||| |||| || || |||||| ||  ||  ||  | | | ||
Db    20 GCCGGCGGGCCAAAGTGACGCCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy   110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
            ||  ||||||||       |||||| ||| |||||||||| |||||||||  |  ||
Db    80 CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134

Qy   170 AGTGGGGTCTGGAGACTATGGACTCCAACAAGGAACCCTGCTTCCGGATGAAAAACGTCCA 229
            | |||| ||  ||||||||||||||| | |||||||||||| || | ||| || | | 
Db   135 AATGGGGCTCAGGGGACTATGGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy   230 TTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTTGACTGGCATTGTCGG 289
            ||||||||  ||||||||||||||||||||| |||||||||||| ||||||||||| ||
Db   195 TTTCAATAAAATCTTCCTGCCCACCATCACTCCATCATCTTCTTAACTGGCATTGTGGG 254

Qy   290 CAATGGATTGGTGATCCTGGTCATGGTCATGGGTTACCAGAAGAAGTAAGGAACATGACGGACAA 349
            ||||||||||||  ||||||||||||||||||||||||||||||||  |   ||||||||||||
Db   255 CAATGGATTGGTCATCCTGGTCATGGTCATGGGTTACCAGAAGAAGAAACTGAGAAGCATGACGGACAA 314

Qy   350 GTACCGGCTGCACCTGTCAGTGGCTGACCTGTCAGTCGTTCATCATCACACTCCCCTTCTCGGC 409
            ||||  | ||||||||||||| | ||||| || ||| |||||||| | ||  ||||||||| ||
Db   315 GTACAGGCTGCACCTGTCAGTGCCGACCTCCTCCTTCTTTGTCATCACGCTTCCCTTCTCGGGC 374

Qy   410 AGTTGATGCCATGGCCATGTGACTGTACTTTGGGAAATTTGTGTAAGGCTGTCCATATCAT 469
            ||||||| | ||||  | ||| ||| ||||||| | |||  | |||||| ||| ||||| ||
Db   375 AGTTGATCCGTGCAAACTGGTACTTTGGGGACTTCCTATGCAAGGCAGTCATGTCAT 434

Qy   470 CTACACTGTCAACCTCTCAGAGCAGGTTCTCATCCTGCCTTCATCCTGCCTCTGAGCCTGACCGGTA 529
            ||||| ||||||||||||||||  ||| | || ||||||||||||||||| ||  |||||||||| |
Db   435 CTACACAGTCAACCTCTCAGAGAGTGTCCTCATCCTGGCCTTCATCCTGGCCTTCATCCTGGCCTCGGACCGGCTA 494

Qy   530 CCTCGCCATTGTCACGCACCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGC 589
            ||| || || || |||| |||||||||||||| |||||||| |  | |||||||||| |
Db   495 CCTGGCCATCGTCACGCCACCAACAGTCAAGGAGGCCAAGGAGAAGTGTGGCTGTTGGCTGAAAAGGT 554

Qy   590 AGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGC 649
            ||||||| | |||||||  ||||| ||||||| ||||| || |||||||||||||||
Db   555 GGTCTATGTTGGGCGTCTCGATCCCGCCCCTCGCTGCTCGACTATTCCCGACTTCATCTTTGC 614

Qy   650 CGACGTCAGCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGAGACCGCCTTA 709
            | ||| |||  |  ||| |||| |  ||| ||||| | ||| |||||||||||||| ||
Db   615 CAACG----------------TCAGTGAGGCAGCATGACAGCAGAGATATATCTGTGAGACCGCCTCTA 659

Qy   710 CCCCGATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAAATGGTGGTCTCATCCT 769
            ||||  ||  |||||| |||||| || |||| |||||| |  | ||||||| ||||||
Db   660 CCCCAATGACTTGTGGGTGGTTGTTGTTCCAGTTCAGCCACACATCATGGTTGGCCTTATCCT 719

Qy   770 GCCCGGCCATCGTCATCCTCCTGTTACTGGTCATCATCATCCTCTAAGCTGTCACACTCCAA 829
            ||| || || ||||| ||| |||| | ||| || ||| | || ||||||||| ||||||||
Db   720 GCCTGGTATTGTCATCCTGTCCTGCTATTGCATTACCATCTCCAAGCTGTCCACACTCCAA 779

Qy   830 GGGCACCAGAAGCGCAAGGCCTCAAGACGAGTCAAGAGGCCTCATCCTCAGCTTCTTTGC 889
            |||| ||||||||||||||| ||||| |  |||| | | | ||||||||||||||||||
Db   780 GGGCCACCAGAAGCGCAAGGCCCTCAAGACCACCAGTCAGTCATCCTCAGCTTCTTTGC 839

Qy   890 CTGCTGGCTGCCATATTATGTGGGATCAGCATCGACTCCTTCATCCTTTTTCATCTTTTGGGAGTCAT 949
```

```
Qy  1430 AAGTTCTTAGTAGTCTGTTTATCTGTGTGTAGGACTGTGAGGACTGTAGGAGGAAGAAACTGA 1489
            ||||||||||  ||||| | |||||| | ||||||||||          |||| |||||||
Db  1362 AAGTTCTTAGTCTGTGTATGTCTCGTGGTAGGACTGTAGAA-------AAGGGAACTGA 1413

Qy  1490 ACATTCCAGAGATGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCA 1549
            |||||||||||| | |||  | ||| | |||||||  ||| | ||  ||||||  |||||
Db  1414 ACATTCCAGAGAGCGGTGAGTCGTGAATCACGTAAAGCTAGAAATGATCCCCAGCTGTTTATGCA 1473

Qy  1550 TAATCTCTTCATTCCGAGGAGCACCCCCACCCCCACCCCCATTCTTAAAT 1609
            ||  ||||| ||
Db  1474 TAGATAATCTCTCCA------------------- 1488

Qy  1610 TGTTTGGTTATGCTGTGTGATGGTTGTTTTTTTTTGTTGTTGTTGTTTTTT 1669
                         |||||| ||||||||||| |||||||||||||||| ||||
Db  1489 -------------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGAT 1525

Qy  1670 TTTTCGTGAAAGATGGCACTTAAAACAAAGCCTGAAATGGTGGTAGAAATGCTGGGGT 1729
            ||| | || | |||||||||||| ||| |||||| ||||| | ||||||||||||---
Db  1526 TTTGCTGTGAAGATGGCACTTATAACCAAAGCCCAAAGTGGT-ATAGAAATGCTGG--- 1581

Qy  1730 TTTTTTTGTTGTTGTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAAT 1789
                         ||||| |||||||| || ||| |  ||| ||| ||||| |||| ||
Db  1582 -------------TTTTTCAGTTTTCAGGAGTGGGGTTGATTTCAGCCACCTAC-AGT 1623

Qy  1790 GTACAGTCTTGTATTACATTGTTAATAAAAGTCAATGATAAACTTAAAAAAAAAAAA 1849
            |||||||||||||||| | ||||||||||||| || | |||||||||||||||||||
Db  1624 GTACAGTCTTGTATTAAGTTGTTAATAAAAGTACATGTTAAACTTAAAAAAAAAAAAA 1683

Qy  1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
            ||||||||||||||||||||||||||||
Db  1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

RESULT 8
US-08-801-228-3
; Sequence 3, Application US/08801228
; Patent No. 5922541
GENERAL INFORMATION:
APPLICANT: Lee, James
APPLICANT: Wood, William I.
TITLE OF INVENTION: METHODS FOR DETECTION AND AMPLIFICATION OF
TITLE OF INVENTION: PF4A RECEPTOR NUCLEIC ACID
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,228
FILING DATE: 19-Feb-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284586
FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2P1D3

APPLICANT: Lee, James
APPLICANT: Wood, William I.
TITLE OF INVENTION: PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Genentech, Inc.
  STREET: 1 DNA Way
  CITY: South San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94080
COMPUTER READABLE FORM:
  MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/104,296
  FILING DATE: 24-June-1998
  CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/701265
  FILING DATE: 22-AUG-1996
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/664228
  FILING DATE: 06-JUN-1996
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/076093
  FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/810782
  FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Love, Richard B.
  REGISTRATION NUMBER: 34,659
  REFERENCE/DOCKET NUMBER: P0706P2C2
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415/225-5530
  TELEFAX: 415/952-9881
  TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
  LENGTH: 1737 base pairs
  TYPE: Nucleic Acid
  STRANDEDNESS: Single
  TOPOLOGY: Linear
US-09-104-296-3

Query Match   51.9%;   Score 974.4;   DB 3;   Length 1737;
Best Local Similarity   76.1%;   Pred. No. 2.3e-195;
Matches 1391;   Conservative   0;   Mismatches 301;   Indels 136;   Gaps 10;

```
Qy   50 GCAGTTAGCAGTGACCCTCTGAGGCGTTTGGTGTCCGGTAACCACCACGGCTGTAGAGC 109
           ||||||||||||||||||||||| |||||||||||||||| |||||||||||||||||
Db   20 GCGCGGCGCGCAAAGTGACGCGCCGAGGGCCTGAGTGCTTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy  110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTTCTGAAGA 169
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   80 CAGCCGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTAACTACACCGAGGA 134

Qy  170 AGTGGGGTCTCGGAGACTATGACTCCAACAAGGAACCCGTTCCGGGATGAAAACGTCCA 229
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  135 AATGGGCTCAGGGACTATGACTCCATGAAGGAACCCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy  230 TTTCAATAGGATCTTCCTGCTGCCCACCATCACTTCATCATCATCTTCTTGACTGCATAGTCGG 289
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  195 TTTCAATAAAATCTTCCTGCCCACCATCACTCACTCCATCATCATCTTCTTAACTGGCATTGTGGG 254

Qy  290 CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACCGGACAA 349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  255 CAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGCATGAAGCATGACCGGACAA 314

Qy  350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCTTCTGGGC 409
```

RESULT 9
US-09-104-296-3
; Sequence 3, Application US/09104296
; Patent No. 6087475
; GENERAL INFORMATION:

```
Db  840 CTGTTGGCTGCCTTACTACATTGGGGATCAGCAGCACTCCTTCATCCTCCTGGAAATCAT 899
           |||||||||||||||||||||||||||||||||||||| |||||||| |||||||
Qy  950 CAAGCAAGGATGACTGACTGCACAAGTGACTCCATCACAGAGGCCCT 1009
           ||||||||||||||||||||||||||||||| ||||||||||||
Db  900 CAAGCAAGGGTGTGAGTTGAGAACACTGTGCACAAGTGGGATTCCATCACCGAGCCCT 959
           ||||||||||||||||||||||||||||||||||||||
Qy 1010 CGCCCTTCTTCCACTGTTGCCTGGAACCCCATCCTCTATGCCTTCCTCGGGGCCAAGTTCAA 1069
           |||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  960 AGCTTTCTTCCACTGTTGTCTGAACCCCATCTCTATGCTTTCCTGGGAGCCAAATTTAA 1019
           ||||||||||||||||||||||||||||||||
Qy 1070 AAGCTCTGCCCAGCATCCACTCAACTCCATGAGCAGGCTCCAGCTTCAAGATCCTTTC 1129
           ||||||| |||||||||||| ||||||||||||||||||||||||
Db 1020 AACCTCTGCCCAGCACGCACTCACCCTCGTGAGCAGGGGTCCAGCCTCAAGATCCTCTC 1079
           |||||||||||||||||||||||||||||||||||||||
Qy 1130 CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCA 1189
           ||||||||||||||||||||||||||||||||||||| ||||||
Db 1080 CAAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTCAGTTTTCA 1139
           ||||||||||||||||||||||||||||||||||||||
Qy 1190 CTCCAGCTAACCCTTATGCAAAGACTTATATATATATATATATATGATAAAGAACTT 1249
           |||||||||| ||||||||||||||||||||||| |||||||
Db 1140 CTCCCAGCTAA-------CACACAGATGTAAAAGACTTTTTTTTTATACGATAAATAAACTTTT 1191
           |||||||||||||||||||||||||||||||||
Qy 1250 TTTTATGTTACACATTTTCAGATAATAAGAGACTGACCCAGTCTTGTACAGTTTTTTTTTT 1309
           ||||||||| ||||||||||||||||||||||||||||||||||
Db 1192 TTTTAAGTTACACATTTTCAGATATATAAAAGACTGACCAATATTGTACAGTTTTTTATTGC 1251
           |||||||||||||||||||||||||||||||||
Qy 1310 TTTTTAATTGACTGTTGGGAGTTTATGTTCCCTAGTTTTTGTGAGGTTTGACTTAATTT 1369
           ||||||||| |||||||||||||||||||||||
Db 1252 TTGTTGGATTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATTGACTTTA 1305
           ||||||||||||||||||||||||||||||||||||||||||||
Qy 1370 ATATAAAATATTGTTTTTTTGTTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGGCC 1429
           |||| ||||||||||||||||||||||||||||||
Db 1306 TTTATA---TAAATTTTTTTTTTGTTTCATATTGATGTGTGCTAGGCAGGACCCTGTGGCC 1361
           |||||||||||||||||||||||||||||||||
Qy 1430 AAGTTCTTAGTAGCTGTTTATCTGTGTGTAGGACTGTAGAGGAAGAAACTGA 1489
           ||||||||||||||||| |||||||||||||||| ||||||
Db 1362 AAGTTCTTAGTTGCTGTGTATGTCTCGTGGTAGGACTGTAGAA------AAGGGAACTGA 1413
           ||||||||||||||||||||||||||||||||||||
Qy 1490 ACATTCCAGAAATGTGTGTAAATTGAATAAAGTCAGCCGTGTAATCCTCAGCTGTTGCTGCA 1549
           |||||||||||||||||| |||||||||||||||||||||||||
Db 1414 ACATTCCAGAGCCGTGTAGTGATCACGTAAGCTAGAAATGATCCCAGCTGTTTATGCA 1473
           |||||||||||||||||||||||||||||||||
Qy 1550 TAATCTCTTCATTCCGAGGAGCACCCACCCCCACCCCCATTCTTAAAT 1609
           |||| |
Db 1474 TAGATATATCTCTCCA--------------------- 1488
           |
Qy 1610 TGTTTGGTTATGCTGTGTGATGGTTGTTGTTTTTTTTGTTTTTGTTGTGTTGTTTTT 1669
           || |||||||||||||||||
Db 1489 ---------------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGAT 1525
           ||||||||||||||||||||||||
Qy 1670 TTTTCTGTAAAAGATGGCACTTAAAACCAAAGCCTGAAATGGTGGTAGAAATGCTGGGGT 1729
           ||||||||| |||||||||||||||||||||| |||||||| ||||||
Db 1526 TTTGCTGTAGAAGATGGCACTTATAACCAAAGCCCAAGTGT-ATAGAAATGCTGG---- 1581
           ||||||||||||||||||||||||
Qy 1730 TTTTTTGTTTGTTGTTTTTTCAGTTTTCAGAGTAGATTGACTTCCAGTCCCTACAAAT 1789
           |||| |||||||||||||| |||||||||||
Db 1582 ----------TTTTCAGTTTCAGGAGTGGGTTGATTTCAGCACCTAC-AGT 1623
           ||||||||||||||||||||||||||||||
Qy 1790 GTACAGTCTTGTGATTACATTGTTAAATAAAAGTCAATGATAAATTAAAAAAAAAAAA 1849
           |||||||||||||||||||||||||||| ||| |||||||||||||
Db 1624 GTACAGTCTTGTGATTACATTGTTAAATAAAAGTACAATGATTAAATTAAAAAAAAAAAAA 1683
           ||||||||||||||||||||||||||||
Qy 1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
           ||||||||||||||||
Db 1684 AAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

```
RESULT 10
PCT-US94-06380-2
; Sequence 2, Application PC/TUS9406380
; GENERAL INFORMATION:
;   APPLICANT: Chuntharapai, Anan
;   APPLICANT: Lee, James
;   APPLICANT: Hebert, Caroline
;   APPLICANT: K. Jin Kim
;   APPLICANT: Genentech, Inc.
;   TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATORY DISRODERS
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;    ADDRESSEE: Genentech, Inc.
;    STREET: 460 Point San Bruno Blvd
;    CITY: South San Francisco
;    STATE: California
;    COUNTRY: USA
;    ZIP: 94080
;   COMPUTER READABLE FORM:
;    MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;    COMPUTER: IBM PC compatible
;    OPERATING SYSTEM: PC-DOS/MS-DOS
;    SOFTWARE: patin (Genentech)
;   CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: PCT/US94/06380
;    FILING DATE:
;    CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER: 08/076093
;    FILING DATE: 11-JUN-1993
;   ATTORNEY/AGENT INFORMATION:
;    NAME: Love, Richard B.
;    REGISTRATION NUMBER: 34,659
;    REFERENCE/DOCKET NUMBER: 706P2P1
;   TELECOMMUNICATION INFORMATION:
;    TELEPHONE: 415/225-5530
;    TELEFAX: 415/952-9881
;    TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;    LENGTH: 1737 bases
;    TYPE: nucleic acid
```

```
;
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US94-06380-2

Query Match      51.9%;   Score 974.4;   DB 5;   Length 1737;
Best Local Similarity   76.1%;   Pred. No. 2.3e-195;
Matches 1391;   Conservative   0;   Mismatches 301;   Indels 136;   Gaps   10;

Qy     50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC 109
          |||||||| |||||||   |||||||| ||||||||||||||   ||||||||  ||||
Db     20 GCGCGGCGCAAAGTGACGCCGAGGCCCTGAGTGCTCCAGTGCCACCGCATCTGGAGAAC 79

Qy    110 GAGTGTGTGCCATGGAACCGATCAGTGAGTAGTATATACACTTCTGATAACTACTCTGAAGA 169
          |||||||||||||||     ||||||||||||||||||||||||||||||||||||
Db     80 CAGCGGTACCATGGA----GGGGATCAGTATATACACTCTCAGATAACTACACCGAGGA 134

Qy    170 AGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAGTCCA 229
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 AATGGGCCCAGGGGACTATGAACTCCATGAAGGAAGACCCTGTTCCGTGAAGAAATGCTAA 194

Qy    230 TTTCAATAGGATCTTCCTGCCCACCATCTCATCATCTTCTTGACTGGCATAGTCGG 289
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 TTTCAATAAAATCTTCCTGCCCACCATCTCCTTCTTCTTGTCATCACGCTTCCCTTCGGGC 254

Qy    290 CAATGGATTGGTGATCTCGGTCATGGGCTCAGTGGTTACCCAGAGAGCATGACGGACAA 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 CAATGGATTGGTCATCTGGTCATGGGTTACCCAGGCAGAGCATGACGGACAA 314

Qy    350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTGGGC 409
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 GTACAGGCTGCACCTGTCAGTGGCTGACTTTGGCAAAGTTCATCACGCTTCCCTTCTGGGC 374

Qy    410 AGTTGATGGCCATGGGTCTGGATCCCAGCCCTCCTGACTATACGTGTAAGGCTGTCCATATCAT 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 AGTTGATGCCGTGGCAAATCCCGTTGGATCTTTGGGAACTTCCGATCAAGGCAGTCCATGTCAT 434

Qy    470 CTACACTGTCAACCTCACAGCAGCAGCGTTCTCATCCTGCGCCTTCATCAGCCTGGACCGGTA 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 CTACACAGCAACCTCACAGCAGCAGCAGTGTCCTCATCCTGCGCCTTCATCAGCCTGGACCGCTA 494

Qy    530 CCTCGCCATTGTCCACGGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGC 589
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    495 CCTGGCCATCGTCCACGGCCACCAACAGTCAGAGGCCAAGGAAGCTGTGGCTGAAAAGGT 554

Qy    590 AGTCTATGTGGGGGTCTGGATCCCAGCCCTCCTGACTATACCGATTCATCCTTTGC 649
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    555 GGTCTATGTTGGCGTCTGGATCCGCTGCCTCCCGCGACTTCATCTTTGC 614

Qy    650 CGACGTCAGCCAGGGGACATCAGTCAGGGGGATGACAGTGACATCTGTGACCGCCTTTA 709
          ||||     |||||||||||||||||||||||||||||||||||||||||||||
Db    615 CAACG-------TCAGTGAGGCAGCAGAGATATCTGTGACCGCTTCTA 659

Qy    710 CCCGATAGCCGTGTGGATGGTGTTCAATTCGAGCATAATGGTGGGTCTCATCCT 769
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 CCCCAATGACTTGTGGGGTGGTGTTCCAGTTCAGCACATCATGGTTGGCCTTATCCT 719

Qy    770 GGCCCACCAGAGAGCCCAAGGGCCCTCAAGACACAGTCATCCTCATCCTCTAGCGTCTTCTTGC 829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 GCCTGGTATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAA 779

Qy    830 GGGCCACCCAGAGAGCGCAAGGGCCCTCAAGACAGCAGTCATCCTCATCCTCCTAGCTTCTTGC 889
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 GGGCCACCCAGAGAGCGCAAGGGCCCTCAAGACCACAGTCATCCTCATCCTTCTTCGC 839

Qy    890 CTGCTGGCTGCCATATTATGTGGGGATCAGCAGCATCGACAATCGCTCCTCTTTGGGAGTCAT 949
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 CTGTTGGCTGCCTTACTACATTGGGGATCAGGATCAGCAATCCTCCTTCATCCTCCTGGAAATCAT 899

Qy    950 CAAGCAAGGATGTGACTTCGAGAGCATTGCACAAGTGACAAGTGGATCTCCATCACAGAGGCCCT 1009
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 CAAGCAAGAGGTGGTGAGTTGAGAACACTGTCGCAACAGTGGAATTCCATCACCACCGGAGCCCT 959

Qy   1010 CGCCTTCTTCACTGTTGCCTGCATGATCCTCCGTGAACCCCATCCTCTCATGCCTCTGGGGCAAGTTCAA 1069
```

```
Db    960 AGCTTTCTTCCCACTGTCTGTCACTCAATGCAACATCCATCCTCTCTATGCCTCTTCCTTGGAGGCCAAATTTAA 1019
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1070 AAGCTGCCAGCCATGCACATGACAGCAGGCGTTCCAGCCTCAGAGACGTCTTC 1129
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 AACCTCGCCCAGCACCGCACTCACCCTGTGAGCAGAGGTCCAGCCTCAAGATCCTCTC 1079

Qy   1130 CAAAGGAAAGCGGGGTGGACATCCTTCCGTCTCCACGAGTCAGAATCCTCAGTTTCA 1189
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 CAAAGGAAAGCGAGGTGGACATTCATCCTTCCACTGAGTCTGAGTCTTCAAGTTTCA 1139

Qy   1190 CTCCAGCTAACCCTATGCAAGACTTATATATAATATAATATATGATAAAGAACTT 1249
          |||||||||     |||||||||||||||||||||||||||
Db   1140 CTCCAGCTAA------CACAGATGTAAAAGACTTTTTTTTATACGATAAAATAACTTT 1191

Qy   1250 TTTATGTTACACATTTCCAGATAAGAGACTGACCAGTCCTTGTACAGTTGTTTTTTT 1309
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1192 TTTAAGTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTTTTATTGC 1251

Qy   1310 TTTTAATTGACGTGTTGGGAGTTATGTTCCTCTCAGTTTTTGTGAGGTTGACTTAATTT 1369
          |||||||||||||||       |||||||||||||||||||||||||
Db   1252 TTGTTGGATTTTGT------CTTGGGTTTCTTTAGTTTTTGTGAAGTTTAATTGACTTA 1305

Qy   1370 ATATAAATATTGTTTTTTGTTTGTTTACTGTGTGTAGGACTGTGAGAGGGAGAACTGA 1429
          |||||       ||||||||||||||||||||||||||||||||||||
Db   1306 TTTATA----TAAAATTTTTTTGTTTCATATTGATGTGTCTCAGGGTAGGACTGTGGCC 1361

Qy   1430 AAGTTCTTAGTAGCTGTTTATCTGTGTGTAGGACTGTGTAGAACTGTGAGAGAAACTGA 1489
          |||||||||||||||||||||||||||||       |||||||||||
Db   1362 AAGTTCTTAGTTGCTGTGTATGTCTCGTGGTAGGAC--------AAGGGACTGA 1413

Qy   1490 ACATTCCAGAATGTGTGGTAAATTGAATAAGCAGCCGTGATCCTCAGTGTTGTTGCTGCA 1549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1414 ACATTCCAGAGCGTGTGAGTGAATCAGTAAAGCTAGAAAATGAATCACGAGAAAAGTTTATGCA 1473

Qy   1550 TAATCTCTTCATTCCGAGGAGCACCCCACCCCCACCCCCACCCCATTCTTAAAT 1609
          |||       |||||||||||
Db   1474 TAGATAATCTCTCCA--------------- 1488

Qy   1610 TGTTTGGTTATGCTGTGTGATGGTTTGTTGGTTTTTTTTTGTGTGTGTTTTTT 1669
          |||||||||||||||||||||||||       |||||||||||||||||
Db   1489 -----------------------------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGAT 1525

Qy   1670 TTTTCTGTAAAAGATGCACTTAAAACCAAAGCTGAAAATGTGGTGGGGT 1729
          ||||||||||||||||||||||||||||||||||||||||
Db   1526 TTTGCTGTGAGAGATGCGCACTTATAACCAAGCCCAAAGTGGT-ATAGAAATGCTCGG--- 1581

Qy   1730 TTTTTTGTTTGTTGTTGTTTTCAGTTTTTCAGTTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAAT 1789
          |||||||||||||||||||||||       |||||||||||||||||||||||
Db   1582 ------------TTTTTCAGTTTTCAGGAGTAGGTGGGTTGATTTCAGCACCTAC-AGT 1623

Qy   1790 GTACAGTCTTGTATTACATTGTTAATAAAAGTCAATGATAAAACTTAAAAAAAAAAAAAAA 1849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1624 GTACAGTCTTGTATTAAGTTGTTAATAAAAGTACATGTTTAATAAACTTAAAAAAAAAAAA 1683

Qy   1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877

Db   1684 AAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-517-605-14

Query Match   50.8%; Score 952.8; DB 4; Length 1679;
Best Local Similarity 80.2%; Pred. No. 7.7e-191;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

Qy   44 GCAGGTGCAGTAGCAGTGACCCTGAGCGTTGGTGCTCCGTAACCACACGGGTG 103
         ||||||||||| || | || |||| | |||||| ||| || ||| ||
Db   12 GCGGCAGCAGGTAGCAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy   104 TAGAGGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
         | ||| ||| ||||| | ||| || ||||||||||||||||||||||
Db   72 GAGAACCAGCCGGTTACCATGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGGGACTATGACTCCAACAAGGAACCTGCTTCCGGGATGAAAA 223
         || | ||||||| || || |||||||||||||||
Db   127 CGAGGAAATGGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAA 186

Qy   224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTCACTTCTTCATCATCTTCTTGACTGGCAT 283
         ||||| ||||||||||||| || ||||||||||||||||||||||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTCCATCATCTTCTTCCATCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAGAAGCTAAGGAGCATGAC 343
         || ||| |||| ||| ||||||||||||||||||||| |||| ||||||||||
Db   247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGAACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTGTCATCACACTCCCCTT 403
         || |||||| || ||||| |||||||| ||| || |||||||||||||||||
Db   307 GGACAAGTACAGGCTGCACCTGTCAGTGGCGCGACCTCCTCTTGTCATCACGCGCTTCCCTT 366

Qy   404 CTGGGCAGTTGATGCCATGGCGTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
         || ||||||||||| |||||||||||||| ||||||| || |||||||||||||||
Db   367 CTGGGCAGTTGATGCCGTGGCAAAACTGGTACTTTGGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCCAACCTTCACAGCAGCGTTCTCATCCTCGGCCTTCATCAGCCTGGA 523
         |||||| |||||||| ||||||||| |||||||||| |||||||||||||||||||||
Db   427 TGTCATCTACACAGTCAACCTTCTCTACAGCAGTGTCCTTCATCCTCGGCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTCGGCCATTGTCCACCACCAGTCAAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
         ||||||||| |||||||||||||||||| |||||| ||||| ||||||||||||||||||
Db   487 CCGCTACCTGGCCATCGTCCACGCACCACCAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy   584 AAAGGCAGTCTATGTGGCGTCTGGATCCCAGCCCCTCCTGACTATACCTGACTTCAT 643
         ||||||||||| ||||||||||| || |||||||||||||||||||||||
Db   547 AAAGGTGGTCTATGTTGGCGTCTGGATCCCGTCGCTCCTGCTGACTATTCCCGACTTCAT 606

Qy   644 CTTTGCCGACGTGCAGCCAGCAGGGGACATCAGTCAGTGACAGTACATCTGTGACCG 703
         ||||||||||| |||||||||||| || ||||||||| |||||
Db   607 CTTTGCCAACG-----------TCAGTGAGGCAGATGACAGATACTCGTGACCG 651

Qy   704 CCTTTACCCCGATAGCCTGTGGATGGGTGTGTTCAATTCCAGCATATAATGGTCGGGTCT 763
         |||||||||| ||||||||||| ||||||||||||||||||||||||||||
Db   652 CTTCTACCCCCAATGACTTGTGGGTGTGGTTCAGTTTTCAGCACACATCATGGTTGGCCT 711

Qy   764 CATCCTGCCCGGCATCGTCATCCTCTCCGTTACTGCATCATCATTTCTCTAAGCTGTCACA 823
         ||||| |||||||||||||||| |||||| | | |||| |||||||||||||||
Db   712 TATCCTGCCTGGTATTGTCATCCTCTCCTGTCCGCTGCCATTATCATCCTCCAAGCTGTCACA 771

Qy   824 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACAGTCATCCTCATCCTCTAGCTTTT 883
         || |||||||||||||||||||||| |||||||||||||||||||||
Db   772 CTCCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACCAGTCATCCTCATCCTCTGGCTTT 831

Qy   884 CTTTGCCTTGCTGGCTGCCATATATGTGGGGATGACTCAGCATCCTTTTGGG 943
         ||||||||||||| ||||| | ||| | |||||||||||||||||||
Db   832 CTTCGCCTTGTTGGCTTGCTGCCTTACTACATTGGGGATCACTACTGACTCCTTGCTGGA 891

Qy 1544 GCTGCATA 1551
         ||||||||
Db 1466 TATGCATA 1473

RESULT 12
US-09-582-224A-5
; Sequence 5, Application US/09582224A
; Patent No. 6429308
; GENERAL INFORMATION:
; APPLICANT: IIJIMA, Osamu
; APPLICANT: GOTO, Takeshi
; APPLICANT: SHIMADA, Takashi
; TITLE OF INVENTION: HIV Infection Inhibitors
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/582,224A
; CURRENT FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: PCT/JP99/06534
; PRIOR FILING DATE: 1999/11/24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cDNA of CXCR4
US-09-582-224A-5

Query Match   50.7%; Score 951.8; DB 4; Length 1664;
Best Local Similarity 80.2%; Pred. No. 1.2e-190;
Matches 1209; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

Qy   944 AGTCATCAAGCAAGGATGTGACTTCGAGAGCCATTGTGCACAAGTGGATCTCCATCACAGA 1003
         || ||||||||||| ||||| |||||| ||| ||||||||||| ||| |||||||||
Db   892 AATCATCAAGCAAGCAAGGGTGTGAGTTTGAGAACTGTGCACACTGTGCACAAGTGGATTTCCATCACCGA 951

Qy 1004 GGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCCTGGGGCCAA 1063
         |||||||| ||||| ||||| |||||||||||||||||||| ||||||||
Db   952 GGGCCCTAGCTTCTTCTCCACTGTTGTCTGAACCCCATCCTCCTATGCTTTCCTTGAGCCAA 1011

Qy 1064 GTTCAAAGCTCTGCCCAGCATGCATCACCAACTCACTCGAGGCAGAGGGCTCCAGAT 1123
         || ||||||||||||||| ||| ||||||| ||||| |||||||||||||
Db 1012 ATTAAAACCTTCTGCCCAGCAGCAGCACTCACCTCGTGAGCAGAAGGGGTCCAGCCTCAAGAT 1071

Qy 1124 CCTTTCCAAAGAAAAGCGGGGGTGGACACTCTTCCCGTCTCCAGAATCCTCCAG 1183
         ||||||||||||||| |||||| ||||||||||| |||||||||||||||
Db 1072 CCTCCCAAAGGAAAGCGAGGTGGACATTCATCGTGTTCCACTGAGTCTTCAAG 1131

Qy 1184 TTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAATATATATATATGATAAA 1243
         ||||||||||||||||||| || |||||| |||||| |||| |||||||
Db 1132 TTTTCACTCCAGCTAACACAGATGTAAAAGACTTTT-------TTTATACGATAAATA 1183

Qy 1244 GAACTTTTTTATGTTACACATTTTCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT 1303
         ||||||||||| ||||||||||| |||||||||||| |||||||||||||||
Db 1184 ACTTTTTTTAAGTTACACATTTTTCAGATATAAAAGAACTGACCAAATATTGTACAGTTTT 1243

Qy 1304 TTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTTGACT 1363
                    |||| ||| ||||||||| |||||||||||
Db 1244 TATTGCTTGTTGGATTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATT 1297

Qy 1364 TAATTTATATAAATATTGTTTTTTTTGTTTTGTTTTCATGTGAATGAGCGTCTAGGCAGGACCT 1423
                  |||||||||| | | |||||||||||||||||||| ||||
Db 1298 GACTTATTTATA----TAAATTTTTTTTGTTTGTTTTCATATTGATGTGTGTCTAGGCAGGACCT 1353

Qy 1424 GTGGCCAAGTCTTTAGTAGCTGTTTATCTGTGTGGTAGGACTGTAGAACTGTAGAGGAGGAGA 1483
         ||||||| |||||||||||| ||||||||||||| |||||||||
Db 1354 GTGGCCAAGTTCTTAGTTGTGCTGTATGTGTCGTGGTAGGACTGTAGAA-------AAGGG 1405

Qy 1484 AACTGAACATTCCAGAATGTGTGTAAATTGAATAAAAGGTGTAAAGCTGTTCCAGCTGTT 1543
         ||||||| |||||||||| ||||||||||||| ||||| |||||||||||
Db 1406 AACTGAACATTCCAGAGCGTGTGAGTGAATCACGTGAGTAGAAATGATCCCCAGCTGTT 1465

Qy 1544 GCTGCATA 1551
         ||||||||
Db 1466 TATGCATA 1473

```
Qy    45  CAGGTGCAGGTAGCAGTGACCCTTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGT  104
          | |||||  |||||| |  ||||| | |||||||||||   ||||||| |||||| ||
Db     1  CGGCAGCAGGTAGCAAAGTGACGCGGAGGGCCTGAGTGCTCCAGTGCTCCAGTGCCACCGGCATCTGG  60

Qy   105  AGACCGAGTGTTGCCATGCAGGATCAGTGTGAGTATATACACTTCTGATAACTACTCT  164
          ||   ||| ||||||||||| ||||-----GGGGATCAGTGTATATACACTTCAGATAACTACCACC
Db    61  AGAACCAGCGGTTACCATGGA-----GGGGATCAGTGTATATACACTTCAGATAACTACCACC  115

Qy   165  GAAGAAGTGGGGTCTGAGACTATGACTCCAACAGGAACCCTGCTTCCGGGATGAAAAC  224
          |||||| ||||| ||||| || ||| |||| |||| |||  ||| ||||| |||||||
Db   116  GAGGAAATGGGGCTCAGGGACTATGACTCCATGAAGGAACCCGTTTCCGTGAAGAAAAT  175

Qy   225  GTCCATTTCAATAGGATCTTCTGCCCACCATTACTTCATCATCTTCTTGACTGGCATA  284
          ||  ||||||||| |||||||||||  ||||||| ||||| |||||||||||| || |
Db   176  GCTAATTTCAATAAAATCTTCCTCGCCCACCATCATCTTCTTCATCATCTTCTTAACTGGCATT  235

Qy   285  GTCGGCAATGGTGGATTGGTCATGGTCATGGGCTCAGTGGCTGAACTCCTTGTCATCACACTCCCCTTC  344
          || ||||||||| ||| |||||  ||| ||||| |||||||| |||| ||| ||||||||||||||
Db   236  GTGGGCAATGGATTGGTCATCCTGGTCATGGGCTTACCAGAAGAAAGTGAGAAGCATGACG  295

Qy   345  GACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTGGCTCAGTGGCTCTTGTCATCACACTCCCCTTC  404
          ||||||||| ||||||||| |||||||| ||| ||||| ||||||| |||  ||||||||
Db   296  GACAAGTACAGGCGTGCACCTGTCAGTGGCCGACCTCCTTTGTCATCACCGCTTCCCTTC  355

Qy   405  TGGGCAGTTGATGCCATGGCCATGGCTGACTGGTGTACTTGGGAAATTTTGTGTAAGGCTGTCCAT  464
          |||||||||||||||||||||||| |||||||  |||||||||||||||  ||||| |||  ||||
Db   356  TGGGCAGTTGATGCCGTGGCAAACTCGGTACTTTGGAACTTCCTATGCAAGGCAGTCCAT  415

Qy   465  ATCATCTACACTGTCAACCTCTACAGCAGCGTTCCATCCTGCCTTCATCAGCCTGAC  524
          |||| |||||||| |||||||||||||||||||  |||||||||||||| |||||||
Db   416  GTCATCTACACAGTCAACCTCTACAGCAGTGTCCCTGGCCTTCATCAGTCTGAC  475

Qy   525  CGGTACCTGCGCCATTGTCCACGGCCACCAACAGTCAAAGGCCAAGGAAAACTGCTGGCTGAA  584
          ||  |||| ||||| ||||||||||| ||||||||||||||||||| ||||||||||||||||
Db   476  CGCTACCTGCCATCGTCCACGGCCACCAACAGTCAGAGGCCAAGGAAGCGTGTTGGCTGAA  535

Qy   585  AAGGCAGTCTATGTGGGGCTCTTGGATCGGGATCCCCAGCCCTCCCTGACTATACCTGACTTCATC  644
          |||||||  ||||||||||| ||||||||| |||  |||||||||||||||||||||||||||||
Db   536  AAGTGGTCTATGTGGCGGGCTCTTGGATCCCCGCTATTCCCGACTATCCCGACTTCATC  595

Qy   645  TTTGCGACGTCAGCCAGCGGGGACATCAGTCAGGGGATGACAGGTACATCTGACCGC  704
          |||||||  | |--------TCAGTGAGGCAGATGACAGATAGAGATATATCTGTGACCGC
Db   596  TTTGCCACG--------TCAGTGAGGCAGATGACAGATAGAGATATATCTGTGACCGC  640

Qy   705  CTTTACCCCGATAGCCTGTGGATGGTGTTTCAATTCCAGCATATAATGTGGGTCTC  764
          ||||||||  ||||  || |||||||| |||| |||||| ||| | |||||||
Db   641  TTCTACCCCAATGACTTGTGGGTGGTGTTGTTGTTCCAGTTCAGCATCATCATGTTGGCCTT  700

Qy   765  ATCCTGCCCGGCATCGTCATCCTCCTGTTACTGCATCATCATCTCTCAAGCTGTCACAC  824
          |||||  |||  ||||| ||||||| |||  || |||  |||| || ||||||||||||
Db   701  ATCCTGGTTGGTTATTGTCATCCTGTCTCCGCTATTGCATTCATCTCCTCAAGCTGTCACAC  760

Qy   825  TCCAAGGGCCACCAGAGAGCGCAAGGCCCTCAGCGACGCCAGCAGTCATCCTCCATCCTCCTAGCCTTTC  884
          ||||||||||| |||||||||||||||||| |||||| ||||||||||||||||||||||||||||  |||
Db   761  TCCAAGGGCCACCAGAAGAGCGCAAGGCCCCTCAGCAGACCACCAGTCATCCTCATCCTCCTGGCTTTC  820

Qy   885  TTTGCCTGGCTGGCTGCCATATTATGTGGGGATCAGCATCGACTCCTCCTTCATCTTTTGGGA  944
          ||||| || |||||| ||| |||||||||||  || ||||| |||||||||||||||||
Db   821  TTCGCCTGTTGGCTGCCTTACTACATTGGGGATCAGCATCGACTCCTCCTTCATCCTCCTGGAA  880

Qy   945  GTCATCAAGCAAGGATGTGACTTCGAGAGCAATTGTGCACAAGTGGATCTTCCATCACAGAG  1004
          ||||||||| ||||||||||||||| |||||||||| |||||||||||  |||||||||||
Db   881  ATCATCAAGCAAGGGTGTGAGTTTGAGAGACTGTGCACAAGTGGATTTCCATCACCCGAG  940

Qy  1005  GCCCTCGCCTTCTTCCACGTTGCCCTGAAGCCCGAACCCATCCTCTATGGCTTCCTCGGGGCCAAG  1064
          || ||||| ||| |||| ||| |||||||||||||||||| ||||| ||||||||||||||| |||
Db   941  GCCCCAGCTTTTCTTCCCACTGTGTCTGGAACCCCCATCCTCTATGTTTCTTTGGAGCCAA  1000

Qy  1065  TTCAAAAGCTCTGCGCCCAGCATGCACCCAGCCACTCAACTCACTCAATGAGCGAGAGGCTTCCAGCCTTCAAGATC  1124
          ||  ||| |||||||||||||  ||| ||||||||||||||| ||||||||||||||||||||| ||||||||||||
Db  1001  TTTAAAACCTCTGCCCCAGCAAGCACGACGACTCACCTCTGTGAGCAGCAGAGGTCCAGCCTTCAAGATC  1060

Qy  1125  CTTTCCAAAGGAAAGGCAGGTGGACACTCTTCCCGTCTCCACGGAGTCAGAATCCTCCAGT  1184
          ||||||||||||||| |||||||||| || ||||  ||||||||||| |||| | ||||||
Db  1061  CTCTCCAAAGGAAAGGCAGGTGGACATTCATCGTTTCCACTGAGTCTGAGTCTTCAAGT  1120

Qy  1185  TTTCACTCCAGCTAACCTTATGCAAAGACTTATATAATAATATATATGATAAAG  1244
          ||||||||||||||||  | ||||| ||||| |            |  ||||||||
Db  1121  TTTCACTCCAGCTAACACAGATGTAAAAGACTTTT------TTTATACGATAAATAA  1172

Qy  1245  AACTTTTTATGTTACACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTT  1304
          | ||||||  ||||| ||||||||| || || |||||| ||| ||||||||||||||||
Db  1173  CTTTTTTTAAGTTACCACATTTTCAGATATAAAAGACTGACCAAATATTGTACAGTTTTT  1232

Qy  1305  TTTTTTTTAATTGACTGTTGGGAGTTTATGTTCCCTAGTTTTTGTGAGGTTTGACTT  1364
          || |  |||| || |||||||| ||     |||||||||||| ||||  ||||| ||
Db  1233  ATTGCTTGTTGGATTTTTGT-----CTTGTGTTCTTAGTTTTTGTGAAGTTTAATTG  1286

Qy  1365  AATTTATATAAATATGTTTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTG  1424
          | ||||| || |   || |||||||||||||||  |||| |||  |||||||||||||
Db  1287  ACTTATTTATA---TAAATTTTTTTGTTTCATATTGATGTGTCTAGGCAGGACCTG  1342

Qy  1425  TGGCCAAGTCTTCTAGTAGCTGTTATCGTGTGTAGGACTGTGAGGAGAAGAA  1484
          |||||||||  ||||| ||||| | || |  ||||||||| | ||      |||||
Db  1343  TGGCCAAGTTCTTAGTTGCTGTATGTCTCGTGGTAGGACTGTAGAA------AAGGGA  1394

Qy  1485  ACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCGTGTTG  1544
          |||||||||||||||| ||||  | ||| | | |||||||  | ||||||| |||| ||||
Db  1395  ACTGAACATTCCAGAGCGTGTAGTTAATCACGTAAAGCTAGAAATGATCCCCAGCTGTTT  1454

Qy  1545  CTGCATA  1551
          | |||||
Db  1455  ATGCATA  1461
```

```
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     LIBRARY: GENBANK
;     CLONE: g2199868
US-09-023-655-1213

Query Match          50.7%;    Score 951.8;  DB 4;  Length 1664;
Best Local Similarity   80.2%;    Pred. No. 1.2e-190;
Matches 1209;  Conservative  0;  Mismatches 252;  Indels 46;  Gaps  6;

Qy     45 CAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTTGGTGCTCCGGTAACCACCAGGCTGT 104
          ||||||| |||||||| |||||||  ||  | | ||  || ||| ||||| ||||| ||
Db      1 CGGCAGCAGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGG 60

Qy    105 AGAGCCAGTGTTGCCATGGAACCGATCAGTGAGTATATACACTTCTGATAACTACTCT 164
          |||||| ||||| |||||||||      ||||||||||||||||||||||||| |
Db     61 AGAACCAGCGGTTACCATGGA----GGGGATCAGTAGTATATACACTTCAGATAACTACACC 115

Qy    165 GAAGAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAC 224
          |||||| |||| |||||||||||  ||||   |  ||||||||||| ||||||||||||
Db    116 GAGGAAATGGGCTCAGGGACGACATGACTCCATGAGAGACCCCTGTTTCCGTGGAGAAAAT 175

Qy    225 GTCCATTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTGACTGGCATA 284
          ||||||||||||||||  |||||||||||||||||||||| ||||||  | ||||||
Db    176 GCTAATTTCAATAAAATCTTCCTGCCACCATCTCCATCATCTTCTTAACTGGCATT 235

Qy    285 GTCGGCAATGGATTGGTGATCCTGGTCATGGTTACCAGAGAGAAGTAAGGACATGACG 344
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    236 GTGGGCAATGGATTGGTCATCCTGGTCATGGTTACCAGAGAGAAACTGAGAAGCATGACG 295

Qy    345 GACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCTCGTTGTCATCACACTCCCCTTC 404
          |||||||| |||||||||||||||||||||||||||||||||||||||  | ||||||
Db    296 GACAAGTACAGGCTGCACCTGTCAGTGGCGCCGACCTCCTCTTTGTCATCACCTTCCCTTC 355

Qy    405 TGGGCAGTTGATGCCATGGCCTGACTGTACTTGGGAAATTTTGTTGAAGGCGTCCAT 464
          ||||||||||| | |||||||||||||||||  ||||||||| |||| ||||||||
Db    356 TGGGCAGTTGATGCCGTGGCCTGACTTTGGGAACTTGTACTTTGGGAACTTCCTATGCAAGGCAGTCCAGT 415

Qy    465 ATCATTCACTGTCAACCTCTACAGCAGCAGCCACCAACAACTGCTGGCTGAA 524
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    416 GTCATCCTCACACAGTCAACCTCTACAGCAGCAGCCACCAACAACTGCTGGCTGGAA 475
```

RESULT 14
US-09-016-434-1235

; Sequence 1235, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

```
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Zeller, Karen J.
;     REGISTRATION NUMBER: 37,071
;     REFERENCE/DOCKET NUMBER: PA-0002 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (650) 855-0555
;       TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1235:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1225 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     LIBRARY: GENBANK
;     CLONE: g189313
US-09-016-434-1235

Query Match             45.8%;   Score 859;   DB 4;   Length 1225;
Best Local Similarity   83.3%;   Pred. No. 3.3e-171;
Matches 1009;   Conservative  0;   Mismatches 180;   Indels  22;   Gaps  2;

Qy    136 TGAGTATATACACTTCGATAACTACCTGAAGAAGTGGGGTCTGAGACTGATGACTCCA   195
          ||||||||||||||||||||||||||  || |||  ||| |||  |||||| ||||||
Db     36 TCAGTATATACACTTCGATAACTACACCGAGGAAATGGGCTCAGGGACTATGACTCCA    95

Qy    196 ACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCA   255
          | |||||||| |||  ||| |||  ||| ||  | |||||||||  |||||||||||||||
Db     96 TGAAGGAACCCGTTCCGTGAAGAAAAATGTCTAATTTCAATAAAATCTTCCTGCCCACCA   155

Qy    256 TCTACTTCATCATCTTCTTGACTGGCATAGTCTGGCAATGATGGTGATCCTGGTCATGG   315
          ||||| |||||||||||||  ||||||| ||  | |||||||||| |||||||||||||
Db    156 TCTACTCCATCATCTTCTTAACTGGCATTGTGGGCAATGATGGTCATCCTGGTCATGG   215

Qy    316 GTTACCAGGAAGAAGCTAAGGAGCAGATGACGGACAAGAGTACCGGCTGCACCTGTCCTG   375
          ||||||||||||||||| ||  | |||||||||||||||| || || |||||||||| ||
Db    216 GTTACCAGGAAGAAGCTGAGAAGACCATGACGGACAAGCAAGTACAGAGGCTGCACCTGTGCCG   275

Qy    376 ACCTCCCTTGTCATCACACTCCCCTTCGGGCAGTGCCAGTGGCCATGGCCTGACTGGTACT   435
          ||||||| ||||||||| ||||||||| |||||||||||||||||   |||| ||||||||
Db    276 ACCTCCCTCTTTGTCATCACGCTCCCCTTCTCGGGCAGTGCCAGTGGCCAAACTGGTACT   335

Qy    436 TTGGGAAATTTTGTGTATAGGCTGTCCATATCATCACACTGTCAACCTCACAGCAGCG   495
          |||||||| ||  |  | ||||| ||||| ||||||| ||||||||||||||||||| |
Db    336 TTGGGAACTTCCTATGCAAGGCAGTCCATGCTCATCAACACAGTCAACCTCACAGCAGTG   395

Qy    496 TTCTCATCCTGGCCTTCATCAGCCTGACCTGGACCGGTACCTGCCATTGTCACGCCACCAACA   555
          | ||||||||||||||||||||||||||||||||||   ||||| | |||||||||||||||
Db    396 TCCTCATCCTGGCCTTCATCAGCCTGACCTGGACCGTCCATGCTCCAGCCACCACCAACA   455

Qy    556 GTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCCGTCTCGGATCCCAG   615
          |||| |||||||||||||| |||||||||||| | ||||||| |||| ||||||||||| ||
Db    456 GTCAGAGGCCAAGGAAACTGTTGGCTGAAAAGTGGTTCTATGTTGGGCGTCTCGGATCCCTG   515

Qy    616 CCCTCCCTGACTATACCCTGACTTCATCTTTGCCGACGTGAGCAGCAGGGGGACATCAGTC   675
          |||||| || ||||| || |||||||||||||| |||             |||| ||||
Db    516 CCCTCCTTGCTGACTATTCCCGACTTCATCTTTGCCAACG----------TCAGTG   560

Qy    676 AGGGGGATGACAGGTACAGTCATCTGTGACCGGCCTTACCCCGATAGCCTGTGGATGGTGGTGT   735
          ||| |||||||||| | ||| ||| |||||||||||||   ||||| | ||||||| ||||||
Db    561 AGGCAGATGACAGCAGATATATCTGTGACCGGCCTTCTCACCCCAATGACTTGTGGGGTGGTGTGT   620

Qy    736 TTCAATTCCAGCATAATGGTGGTCATCCTGCCCGGCCATGTCATCCTTCCTGTT   795
          | ||| |||||||  ||  |||| || ||||| ||| |||||||||||| |||
Db    621 TCCAGTTTCAGCACCATCATGTTGGCCTTATTGTCATCCTGCCTGCT   680

Qy    796 ACTGCATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAGAGCGCAAGGCCCTCA   855
          | ||||| ||||||||   |||||||||||||||||||||||||||| |||||||||||||
Db    681 ATTGCATTATCATCTCCCAAGCTGTCACACTCCAAGGGCCACCAGAGAAGCGCAAGGCCCCTCA   740

Qy    856 AGACGACAGTCATCCATCCTGCCTC...   915
          ||||||||||||||||||||||||
```

```
Db    741 AGACCACAGTCATCCTCATCCTGCTTCTTCGCCTGTTCCTCGTGCTGGCTGGCCTTACTACATTGGGA   800

Qy    916 TCAGGATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGGAGAGCA   975
          ||| ||||||||||||||||| || ||| ||||||||||| ||| |||||  | ||||| |
Db    801 TCAGCATCGACTCCTTCATCCTCCTGGAAATCATCAGCAAGGTGTGAGTTTGAGAACA   860

Qy    976 TTGTGCACAAGTGGATCTTCATCACAGAGGCCCTCGCCTTCTTCCACTGTGCCTGAACC   1035
          |||||||||||||||||| || ||| |||||| | |||||||||||||| | |||||||
Db    861 CTGTGCACAAGTGGATTTCCATCACCGAGGCCCCAGCTTTCTTCCACTGTTGTCTGAACC   920

Qy   1036 CCATCCTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGTCTGCCCAGCATGCACTCAACT   1095
          ||||||||||||| |||||| ||  || |   ||||  |||||||||| ||||| |||||
Db    921 CCATCCTCTATGCTTTCCTTGGAGCAAATTAAAACCTCTGCCCAGCAGCACTCACCT   980

Qy   1096 CCATGCAGAGGCCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGGTGGGACACTCTT   1155
          | ||| |||||| |||||||||||||||| ||||| ||||||| | ||||| || ||
Db    981 CTGTGAGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAGAAAGCGAGGTGGACATTCAT   1040

Qy   1156 CCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCACGCTAACCCTTATGCAAAGACT   1215
          | ||| ||||| |||||| |||  ||||||||| ||| |||||||| |  ||| ||||
Db   1041 CTGTTTCCACTGAGTCAGTCTTCAAGTTTCAACTCCACGCTAACACAGATGTAAA----   1096

Qy   1216 TATATAATAATATATATATATATAGAGAACTTTTTATGTTACACATTTTCAGATAT   1275
            || || ||| |||||||||| |||  |  ||||||| ||||||||||||||||||
Db   1097 --AGACTTTTTTTTTATACGATAAAATAACTTTTTTTTAAGTTACACATTTTCAGATAT   1153

Qy   1276 AAGAGACTGACCAGCGTCTTGTACAGTTTTTTTTTTTTTTTTATTGACTGTTGGGAGTTTAT   1335
          |||||||||||||   | |||||||||||    |||||||||| | |  |||| ||||||
Db   1154 AAAAGACTGACCAATATTGTACAGTTTTTATTGCTTGTTGTGATTTTTGTCTTGTGTTTCT   1213

Qy   1336 GTTCCTCTAGT   1346
            |||  |||
Db   1214 TTAGTTTTTGT   1224
```

```
;    INFORMATION FOR SEQ ID NO:  45:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  1317 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;      MOLECULE TYPE:  cDNA
;      FEATURE:
;        NAME/KEY:  CDS
;        LOCATION:  201..1211
US-08-153-848-45

Query Match        44.5%;   Score 834.6;  DB 1;  Length 1317;
Best Local Similarity   83.9%;   Pred. No. 4.5e-166;
Matches 977;  Conservative   0;  Mismatches 164;  Indels   24;  Gaps   2;
```

```
Qy     134 TGTGAGTATATACACTTCTGATAACTACTTCGAAGAGTGGGTCTGAGACTATGACTC  193
               |||||| || ||||||||||||||| ||  |  ||||||||| ||||| |||||||||||
Db     164 TTTGCAGATATACACTTCAGATAACTACACACCGAGGAAATGGGCTCAGGGACTATGACTC  223

Qy     194 CAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCAC  253
               ||||||| |||||||||||| |||||||||||||||||| |||||||||||
Db     224 CATGAAGGAACCCTGTTCCGTGAAGAAATGCTAATTTCAATAAAATCTTCCTGCCCAC  283

Qy     254 CATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCAT  313
               |||||||| ||||||||||||| |||||||||| |||||||||||||||||||||||||||
Db     284 CATCTACTTCCATCATCTTCTTAACTGCATTGTGGGCAATGGATTGGTCATCCTGGTCAT  343

Qy     314 GGGTTACCAGAAGAAGAGCTAAGGAGCATGAGCGACGACAAGTACCGGCTGCACCTGTCAGTGGC  373
               ||||||| ||||| || ||||| ||||||||||||||||||||||||||||||||||
Db     344 GGGTTACCAGAAGAAAACTGAGAAGCATGAGCAAGTACAGGCTGCACCTGTCAGTGGC  403

Qy     374 TGACCTCCCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTA  433
               ||||| |||||||||||||||||| |||||||||||||||||||| ||||| |||||||
Db     404 CGACCTCCCTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCGTGGCAAACTGGTA  463

Qy     434 CTTTGGGAAATTTTGTGTTAAGGCTGTCCATATACATCTACACTGGTGATCGGTCATCAG  493
               ||||||||| |||||| ||||||||||||||||||||||||||||||||||| |||||||
Db     464 CTTTGGGAACTTCCATCCTGGCCTTCATCAGTCTGGACCTCATCGTCCACGCCACCAA  523

Qy     494 CGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCACGCCACCAA  553
               ||||||||||||||||||| ||||||||||||||||||||| |||||||||||||||||
Db     524 TGTCCTTCATCCTGGCCTTCATCAGTCTGGACCGTCGGACCTTGGCCATCTTTGCCACG  583

Qy     554 CAGTCAAAGGCCAAGGAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCC  613
               ||||||||||| ||||||||||| ||||||||||||||| |||||||||||||||||||
Db     584 CAGTCAGAGGGCCAAGGAAGCTGTTGCTGGAACAGTTGGTTATGTTGGGCTCTGGATCCC  643

Qy     614 AGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGCTCAGCGCCAGGGGACATCAG  673
               ||||||||||||||||||||||||||||| |||||||||||
Db     644 TGCCCCTCCTGCTGACTATTCCCGACTTCATCTTTGCCACG------------TCAG  688

Qy     674 TCAGGGGGATGACAGTGCATCCTCATCCTGTGACCCCTTACCCGATAGCCTGTGGATGGTGGT  733
               ||||||||||||||| ||||||||||| ||||||||||||||||||||| |||||||||
Db     689 TGAGGCAGATGACAGTCATCCTCATCCTGTGACCCCAATGACTTGTGGTGGTGTTGT  748

Qy     734 GTTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCTCTCCTG  793
               ||||||| |||||||| |||||||||||||||||| ||||||||||||||||||||||
Db     749 GTTTCCAGTTTCAGCGATGCACACATCATCGTGGCCTTATTGTCATCCTGCTCCTG  808

Qy     794 TTACTGCATCATCATCCTAAGCTGTCACACTCCAACTCCAAGAGGGCCACCAGAAGCGCAAGGCCCT  853
               ||||||||| ||| ||| ||||| |||||||||||||||||||||||||||||
Db     809 CTATTGCATTATCATCCCAAGCTGTCACACTCCAAGAGGCCACCAGAAGCGCAAGGCCCT  868

Qy     854 CAAGACGACAGTCATCCTCATCCTGCAGCTTCTTTGCCCTGCTGCCATATTATGTGGG  913
               ||||||| ||||||||||||| ||||||||||||| ||||| |||||||||||
Db     869 CAAGACCACAGTCATCCTCATCCTGGCTTCTTTCCGCCTTGCCTTACTACATTGG  928

Qy     914 GATCAGCATGACTGACTCCTTCATCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAG  973
               |||||||||| |||||||||||||| || || ||| |||||||||||||||||||||
Db     929 GATCAGCATCGACTGACTCCTTCATCTTCCTGGAAATCATCAAGCAAGCAAGGGTGTGAGTTTGAGAA  988

Qy     974 CATTGTGCACAAGTGGATCTCCATCACAGAGAGGCCCGCCTTCTTCTTCCACTGTTGCCTGAA  1033
```

```
Db     989 CACTGTGCACAAGTGGATTTCATCACCGAGGCCCCAGCTTTCTTCCTCCACTGTTGCTGAA  1048
               || ||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    1034 CCCATCCTCTATGCCTTCCTGCGGGGCCAAGTCAAAAGTCTGCCCAGCATGCACTCAA  1093
               ||||||||||||||||||||||| ||||||||||||||||| |||||||||||||||
Db    1049 CCCATCCTCTATGCCTTCCTTGGAGCCAAATTAAAACCTCTGCCCAGCACGCACTCAC  1108

Qy    1094 CTCCATGAGCAGAGGCTCCAGCCTCAGATCCTTCAAAGAGAAAGCGGGGTGGACACTC  1153
               ||| |||||||||||||||||||||||| |||||||||||||||| |||||||||||||
Db    1109 CTCTGTGACGAGAGGGTCCAGCCTCAGATCCTCCAAAGGAAGCGAGGTCGGACATTC  1168

Qy    1154 TTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTAACCCTTATGCAAAGA  1213
               |||   ||| ||| ||| |||||| ||||||||||||| ||||| |||
Db    1169 ATCTGTTTCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAACACAGATGTAAA--  1226

Qy    1214 CTTATATAATAATATATATATATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGAT  1273
                                      ||||||||||||||||||||||||||||||||||||
Db    1227 ------AGACTTTTTTTATACGATAAATAACCTTTTTTTAAGTTACACATTTTTTCAGAT  1279

Qy    1274 ATAAGAGACTGACCAGTCTTGTACA  1298
               |||||||||||||| ||| |||
Db    1280 ATAAAAGACTGACCAATATTGAAAA  1304
```

Search completed: May 17, 2004, 16:15:39
Job time : 171.281 secs

```
OM nucleic - nucleic search, using sw model

          Copyright   GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:   May 17, 2004, 12:11:00 ; Search time 918.766 Seconds
                                   (without alignments)
                              9271.041 Million cell updates/sec

Title:           US-09-367-052-1
Perfect score:   1877
Sequence:        1 ccatcctaatacgactcact........aaaaaaaaaaaaaaaaaa 1877

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2947324 seqs, 269024515 residues
Total number of hits satisfying chosen parameters:        5894648

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
            1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
            2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
            3:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
            4:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
            5:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
            6:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
            7:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
            8:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
            9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
           10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
           11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
           12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
           13:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
           14:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
           15:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
           16:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
           17:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
           18:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
           19:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	974.4	51.9	1737	9	US-09-104-063-3	Sequence 3, Appli
2	974.4	51.9	1737	13	US-10-666-689-3	Sequence 3, Appli
3	952.8	50.8	1679	13	US-10-211-462-80	Sequence 80, Appl
4	952.8	50.8	1679	13	US-10-181-906-9	Sequence 9, Appli
5	952.8	50.8	1679	13	US-10-342-887-912	Sequence 912, App
6	952.8	50.8	1679	13	US-10-151-274-14	Sequence 14, Appl
7	952.8	50.8	1679	13	US-10-172-118-912	Sequence 912, App
8	952.8	50.8	1679	13	US-10-170-385-332	Sequence 332, App
9	952.8	50.8	1679	15	US-10-225-557A-75	Sequence 75, Appl
10	952.8	50.8	1679	15	US-10-021-660-58	Sequence 58, Appl
11	952.8	50.8	1679	16	US-10-341-434-177	Sequence 177, App
12	952.8	50.8	1711	10	US-09-971-392-20	Sequence 20, Appl
13	952.8	50.8	1711	15	US-10-101-510-459	Sequence 459, App
14	951.8	50.7	1664	17	US-10-641-643-1213	Sequence 1213, Ap
15	944.8	50.3	1670	9	US-09-880-107-2143	Sequence 2143, Ap
16	944.8	50.3	1670	10	US-09-960-706-636	Sequence 636, App
17	944.8	50.3	1670	16	US-10-372-683-1	Sequence 1, Appli
18	944.8	50.3	1670	16	US-10-440-464-77	Sequence 77, Appl
19	934.4	49.8	5161	16	US-10-160-401-1	Sequence 1, Appli
20	859	45.8	1225	15	US-10-101-510-741	Sequence 741, App
21	859	45.8	1225	16	US-10-305-720-1235	Sequence 1235, Ap
22	847	45.1	1224	15	US-10-101-510-674	Sequence 674, App
23	809.6	43.1	1059	13	US-09-813-651B-84	Sequence 84, Appl
24	809.6	43.1	1059	15	US-10-014-322A-125	Sequence 125, App
25	809.6	43.1	1059	16	US-10-160-401-2	Sequence 2, Appli
26	808	43.0	1102	9	US-09-870-759-143	Sequence 143, App
27	808	43.0	1102	10	US-09-751-708A-143	Sequence 143, App
28	804.8	42.9	1059	11	US-09-826-509-484	Sequence 484, App
29	576.4	30.7	1902	9	US-09-953-692-1	Sequence 1, Appli
30	576.4	30.7	1902	9	US-09-953-717-1	Sequence 1, Appli
31	530	28.2	1076	15	US-10-017-161-789	Sequence 789, App
32	530	28.2	1076	16	US-10-292-798-687	Sequence 687, App
33	345.2	18.4	507	13	US-10-085-783A-14955	Sequence 14955, A
34	345.2	18.4	507	16	US-10-242-535A-14955	Sequence 14955, A
35	303.2	16.2	421	9	US-09-796-692-2611	Sequence 2611, Ap
36	303.2	16.2	421	15	US-10-040-862-2611	Sequence 2611, Ap
37	303.2	16.2	421	16	US-10-057-475B-2611	Sequence 2611, Ap
38	303.2	16.2	421	16	US-10-154-884B-2611	Sequence 2611, Ap
39	287	15.3	487	9	US-09-796-692-8936	Sequence 8936, Ap
40	287	15.3	487	15	US-10-040-862-8936	Sequence 8936, Ap
41	287	15.3	487	16	US-10-057-475B-8936	Sequence 8936, Ap
42	287	15.3	487	16	US-10-154-884B-8936	Sequence 8936, Ap
43	198.4	10.6	1119	12	US-10-997-722-162	Sequence 162, Appl
44	198.4	10.6	1119	15	US-10-251-385-65	Sequence 65, Appl
45	198.4	10.6	1679	9	US-09-104-063-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-09-104-063-3
; Sequence 3, Application US/09104063
; Patent No. US2002016836A1
; GENERAL INFORMATION:
;  APPLICANT: Lee, James
;  APPLICANT: Wood, William I.
;  TITLE OF INVENTION: PF4A Receptors
;  NUMBER OF SEQUENCES: 6
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Genentech, Inc.
;  STREET: 1 DNA Way
;  CITY: South San Francisco
;  STATE: California
;  COUNTRY: USA
;  ZIP: 94080
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: WinPatin (Genentech)
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/09/104,063
;  FILING DATE: 24-June-1998
;  CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: 08/701265
;  FILING DATE: 22-AUG-1996
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: 08/664228
;  FILING DATE: 06-JUN-1996
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: 08/076093
;  FILING DATE: 11-JUN-1993
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: 07/810782
;  FILING DATE: 19-DEC-1991
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Love, Richard B.
;      REGISTRATION NUMBER: 34,659
;      REFERENCE/DOCKET NUMBER: P0706P2C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 415/225-5530
;      TELEFAX:   415/952-9881
;      TELEX:     910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 1737 base pairs
;      TYPE: Nucleic Acid
;      STRANDEDNESS: Single
;      TOPOLOGY: Linear
US-09-104-063-3

Query Match        51.9%;  Score 974.4;  DB 9;   Length 1737;
Best Local Similarity 76.1%;  Pred. No. 5.3e-214;
Matches 1391;  Conservative 0;  Mismatches 0;  Indels 136;  Gaps 10;
```

```
Qy    50 GCAGGTAGCAGTGACCCTCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC 109
         |||||||||||||||| ||||||||| |||||||   ||| ||||||||||||| ||||||
Db    20 GCGCGGGCGCGAAAGTGACGCGCGAGGCCCTGAGTGCTGTTCCAGTAGCCACCGCATCTGAGAAC  79

Qy   110 GAGTGTTGGCCATGGAACCGATCAGTGTGAGTAATAAACTTCTGATAACTACCTCTGAAGA 169
         |||||||||| |  ||||    |||||||||     |||||||||||||  ||||||||||
Db    80 CAGCGGTTACCCATGGA------GGGGATACGTCAGTATATACACTTCAGATAACTACACCCAGGA  134

Qy   170 AGTGGGGTCTGGAGACTATGACTTCCAACAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
         |||||| ||||||||||||||  |||||||||||||||| |||||||||||||||||||
Db   135 AATGGGCTCAGGGGACTATGACTTCCATGAAGGAACCCGTGTTTCCGTGAAGAAAATGCTAA  194

Qy   230 TTTCAATAGGATTTCTCGCCCACCATCTTCTGAACTCTTTGACTGGCATAGTAGTCGG 289
         ||||||||||||||||||| |||||||||||||||||  |||||||||| ||||||||
Db   195 TTTCAATAAAATCTTCCTGCCCACCATCTTCTTCTTACTCCATCATCTTCTTAACTGGCATTGTGTGGG  254

Qy   290 CAATGGATTGGTGATCCTCGTCATGGGTTACCAGGAAGCTAAGGAGCATGACGGACAA 349
         |||||||||||||||||||||||||||  |||||||||| ||||||||||||||||||
Db   255 CAATGGATTGGTCATCCTCGTCATGGGTTACCAGGTTACCAGAGAAGAAGAAACTGAGAGAACATGACGGACAA  314

Qy   350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCCTTCCCCCTTCTGGGC 409
         |||||||||||||||||||||||||||||||||||||  || |||||||||
Db   315 GTACAGGCTGCACCTGTCAGTGGCTGCGCCGACCTCCTCTTTGTCATCACGCTTCATCACGCTTCCTTCTCTGGGC  374

Qy   410 AGTGATGCCATGGCCAGGGGGGACATCATGTGTAAGGCTGTCCATATCAT 469
         ||||||||||||||  ||||| ||||| |||||||||||| ||| |||||
Db   375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGGAACTTGGGACTTTGGAACCTTGATGTCCATGTCAT  434

Qy   470 CTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTCATCGCCTGACCGGTA 529
         ||||||||||||||||||||||||| |||||| |||| ||||| |||||||
Db   435 CTACACAGTCAACCTCTACAGCAGTGTCCTCATCTGCTTCATCAGTCTGGACCGCTA  494

Qy   530 CCTGCCATTGTCCACGCCACCAACAGTCAAAGCCAAGGAAACTGCTGGCTGAAAAGGC 589
         ||||||| |||||||||||||||||||||| |||||||||||||  |||||||||||
Db   495 CCTGCCATCGTCCACGCCACCAACAGTCAGAGCCACCGCCAGGCCGTGTTGGCTGAAAACGT  554

Qy   590 AGTCTATGTGGGCGTCTGGATCCCTGAGTCCAGCCCTCCTCCTGACTATACCTTTGC 649
         ||||||||||||||  |||||| |||||||||||||||| || ||||||||||||||
Db   555 GGTCTCATGTGTGTGGCGTCTGGATCCCCGTCCCCTCCCCTGCTAATTCCCGACTTCATCTTTGC  614

Qy   650 CGACGTCAGCCAGGGGACATCAGTACGGGGATGACAGGTACATCTGTGACCGCCTTTA 709
         ||||||||| |||||||||||| ||||||||||||||||||||||   ||||||||||
Db   615 CAACG------------TCAGTGAGGCAGATGACAGCGGGTTTGATAATGTTGACCTGATAATCCGCCTTCTA  659

Qy   710 CCCCGATAGCCTGTGGATGGTGGTTGTTCAATTCCAGCAGTAATGAATGTGGTCTCATCCT 769
         ||||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db   660 CCCAATGACTTGTGCGGGTGTTGTTGTTCCAGTTTCAGGTTTCAGCCACATCACGGTTGGCCTTATCCT  719

Qy   770 GCCCGGCATCGTCATCCTCCTCCTGTTACTGCATCATCATCTCTAAGCTGTCACACTCCAA 829
         |||||||||||||||||||||||||||||||||||| ||||||||||| |||||||||||
Db   720 GCCTGGTATTGTCATCCTGTCCTGCTATTGCATTATCATCATCTCCAAGCTGTCACACTCCAA  779

Qy   830 GGGGCACCAGGAAGCGCAAGGAGCCCTCAAGACGACAGTCCCTCAGAGACCATCATCCTCATCCTGGCTTCTTCGC 889
```

```
Db   780 GGGGCACCAGGAAGCGCAAGGAGCCCTCAAGACGACAGTCCCTCAGAGACCATCATCCTCATCCTGGCTTCTTCGC 839
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   890 CTGCTGCTGCCATATTATGTGGGATCAGCATCAGCATCGGGATCATCCTCATCCTTTTGGGAGTCAT 949
         |||||||||| ||||||||||||||||||||| |||||||||||||||||||||
Db   840 CTGTTGCTGCCTTACTACATTGGGATCAGCATCAGCATCTTCATCATCCTCCTGAAATCAT  899
Qy   950 CAAGCAAGGATGGTGACTCGAGAGCATTGTGCACAGTGGATCTCCATCACAGAGGCCCT 1009
         |||||||||||||||||||||||||||| |||| |||||||||||||||||||||||||
Db   900 CAAGCAAGGGTGTGAGTTGAGAACAACTGTGTCACAACGTGGATTTCATCACCGAGCCCCT  959
Qy  1010 CGCCTTCTTCCACTGTTGCTGCTGAACCCCATCCTCCTGATGCCTTCCTCGGGGCCAAGTTCAA 1069
         |||||||||||||||| ||||||||||||| ||||||||||| |||||||||||||||||||
Db   960 AGCTTTCTTCCACTGTTCTCTGTGAACCCCATCCTCCTATGCTTCTTCCTTGGAGCCAAATTTAA 1019
Qy  1070 AAGCTGTGCCAGCATGCACTCAACTCCATGAGCAGAGGACAGAGGACGAGGTCCAGCCTCAAGATCCTTTC 1129
         |||||||||||||||||||||  |||||| ||||||||||||||||||||||||||||||||||||||||
Db  1020 AACCTGTGCCAGCACGCACTCACCTCCTGTGAGCAGCAGGGTCCAGCCTCAAGATCCTCTC 1079
Qy  1130 CAAAGGAAAGCCGGGGTGGACACTCTCCGTCTCCAGGAGTCAGAATCCTCCAGTTTTCA 1189
         |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  1080 CAAAGGAAAGCCGAGGTGGACATTCACATTCATCTGTTCCACTGAGTCTGAGTCTTCAAGTTTTCA 1139
Qy  1190 CTCCAGCTAACCCTATGCAAGACTTATAATAATATATATATGATAAAGAACTT 1249
         ||||| ||          ||||||||||||||||||||    ||||||||||
Db  1140 CTCCCAGCTAA------CACAGATGTAAAAGACTTTTTTTATACGATAAATAACTTTT 1191
Qy  1250 TTTATGTTACACATTTCCAGATAAGAAGGACCTTGACCAGTCTTGTACAGTTTTTTTTTT 1309
         |||||||||||||| |||||||||||||||| |||||||||||||||    |||||||
Db  1192 TTTTAAGTTACACATTTTCAGATAAAAGAACTGACCAATATTGTACAGTTTTTAATTGC 1251
Qy  1310 TTTTAATTGACTGTTGGGAGTTTATGTTCCTCTCAGTTTTGTGAGGTTGACTTAATTT 1369
         |||||||||||||  ||||||||||||       |||||||||||||| ||||||||
Db  1252 TTGTTGGATTTTTGT-----CTTGTGTTTCTTTAGTTTTGTGAAGTTTAATTGACTTA 1305
Qy  1370 ATATAAATATTGTTTTTTGTTCATGTGAATGAGCGTCTAGGCAGCAGACCTGTGGCC 1429
         |||          |||||||||||| ||||||||||||||||||||||||||||||
Db  1306 TTTATA----TAAATTTTTTTGTTTCATATTGATGTGTGCTGGAGGACTGTGGGC  1361
Qy  1430 AAGTTCTTAGTAGCTGTTGTTATCGTGTGGTAGGAACTGTGAGGACTGTGAGAGGAGAAAACTGA 1489
         ||||||||||||||||| |||||||||||||||||||||||||||||||       ||||||||||
Db  1362 AAGTTCTTAGTTGCTGTATAGTGTCTCGTGGTAGGAGGACTGTGAGA------AAGGGAACTGA 1413
Qy  1490 ACATTCCAGAATGTGGTAATTGAATAAAGCTAGCCGTGATCCTCAGTGTTGCTGCA 1549
         ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  1414 ACATTCCAGAGCGTGGTAGTGAATCACGTAAAGCTAGAAATGATCCCCAGCGTGTTATGCA 1473
Qy  1550 TAATCTCTTCATTCCGAGGAGCACCCCCCACCCCCACCCCCCATTCTTAAAT 1609
         ||||||||||||||||      |||||||||||||||||||||||||||
Db  1474 TAGATAATCTCCCA---------------------------------------- 1488
Qy  1610 TGTTTGGTTATCGTGTGATGGTTTGTTGGTTTGTTGTTGTTGTTTTT 1669
         ||||||||||||||||||||         |||||||||||||||||
Db  1489 ------------------------TTCCCGTGGAACGTTTTTCCTGTTTCTTAAGACGTGAT 1525
Qy  1670 TTTTCTGTAAAAGATGGCACTTAAAACCAAAGCCTGAAATGCTGGGGT 1729
         |||||||||||||||||||||||||||||||||||||||    |||||
Db  1526 TTTGCTTGGTAGAAGATGGCACTTTATAACCAAAGCCCAAAGTGGT-ATAGAAATGCTGG--- 1581
Qy  1730 TTTTTTGTTTGTTGTTTTTGTTTTCAGTTTTTCAAGAGTAGACTTCAGTCCTACAAAT 1789
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1582 ---------------TTTTTCAGTTTTCAGGAGTTTGATTTGATTTCAGCACCTAC-AGT 1623
Qy  1790 GTACAGTCTTGTATTACATTGTTAATCAATGATAAACTTAAAAAAAAAAAAA 1849
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1624 GTACAGTCTTGTTTAAGTTGTTAATAAAGTACATGTTAAACTGTTAAACTTTAAAAAAAAAAAAAAAA 1683
Qy  1850 AAAAAAAAAAAAAAAAAAAAAAAAA 1877
         ||||||||||||||||||||||||
Db  1684 AAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

RESULT 2

```
US-10-666-689-3
; Sequence 3, Application US/10666689
; Publication No. US2004003783OA1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Human PF4A Receptors, Nucleic Acid Encoding and
; TITLE OF INVENTION: Antibodies Binding Thereto
; FILE REFERENCE: P0706P2C2D2C1
; CURRENT APPLICATION NUMBER: US/10/666,689
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 09/104,063
; PRIOR FILING DATE: 1988-06-24
; PRIOR APPLICATION NUMBER: US 08/701,265
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/664,228
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/076,093
; PRIOR FILING DATE: 1993-06-11
; PRIOR APPLICATION NUMBER: US 07/810,782
; PRIOR FILING DATE: 1991-12-19
; PRIOR APPLICATION NUMBER: US 07/677,211
; PRIOR FILING DATE: 1991-03-29
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
;   LENGTH: 1737
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-666-689-3

Query Match        51.9%;   Score 974.4;  DB 13;  Length 1737;
Best Local Similarity  76.1%;   Pred. No. 5.3e-214;
Matches 1391;  Conservative  0;  Mismatches  301;  Indels  136;  Gaps  10;
```

```
Qy     50  GCAGGTAGCAGTGACCCTCTGAGGCGGTTGGRGCTCGGTAACCACCACGGCTGTAGAGC 109
               ||||||||   | |||||| ||| |||||||||| |||||||
Db     20  -GCGCGGCGCAAAGTGACGCGCGAGGGCCTGAGTGCTTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy    110  GAGTGTTGCCATGGACCGATCAGTGTGAGTATACACTTCTGATAACTACTCCGAGGA 169
               | |||||||   |||| |||||||||||||||||||||||||||
Db     80  CAGCGGTTACCATGA-----GGGGATCAGTATATACATTCAGATAACTACACCGAGGA 134

Qy    170  AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCGCTCCGGGATGAAAACGTCCA 229
               |||||||||||||||||| |||||||||||||||||||  |||||
Db    135  AATGGGCTCAGGGACTATGACTCCATGAGAAGGAACCCATGTTCCCTGAAGAAAATGCTAA 194

Qy    230  TTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTTGACTGCATAGTCGG 289
               |||||||||||||||||||||||||||||||||||||||||||
Db    195  TTTCAATAAAAATCTTCCTGCCCACCATCACTTCATCATCTTCTTAACTGGCATTGTGGGG 254

Qy    290  CAATGGATTGGTGATCCTGGTCATGGTTACCAGAAGAAGCATGAGCATGACGGACAA 349
               |||||||||||||||||||||||||||||||||||  || |  ||||||
Db    255  CAATGGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGCATGAGCATGACGGACAA 314

Qy    350  GTACCGGCTGCACCTGTCAGTGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGTT 409
               |||||||||||||||||||||||||||||||||||||||||
Db    315  GTACAGGCTGCACCTGTCAGTGGCCGACCTTCCTTTGGGAAATTTTGTGTCATCATCGGC 374

Qy    410  AGTTGATGCCATGGCTGACTGGTACTTTGGAACTTCTGGAACTTCCATATCAT 469
               |||||||||||||||||||||||||||||||||||||||
Db    375  AGTTGATGCCGTGGCAAACTGGTACTTTGGAACTTCCATGCAGCAGTCATGTCAT 434

Qy    470  CTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCCTGGAGCCGGTA 529
               ||||||||||||||||||||||||||||||||||||||||||||||
Db    435  CTACACAGTCAACCTCACAGCAGTGTCCTCACAGCCTTCATCCTGGCCTTCATCCTGGAGCCGCTA 494

Qy    530  CCTCGCCATTGTCCACGGCACCGCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGC 589
               |||||||||||||||||||||||
Db    495  CCTGGCCATCGTCACGCACCGCACCAACCACAGTCAGAGGCCAAGGAGCTGGCTGTTGGCTGAAAAGGT 554

Qy    590  AGTCTATGTGGGCGCGTCTGGGATCCCGCCCCTCCCGCTGACTTCATCTTTGC 649
               ||||||||||||||||||||||||||||||||||||||||||||
Db    555  GGTCTATGTTGGCGGTCTGGGATCCCGCCCCTCCCGCTGACTATTCCCGACTCTTGC 614
```

```
Qy    650  CGACGTCAGCCAGCCGGACATCAGTCAGGGGGACATGCAGGTACATCTGTGACCGCCTTA 709
               |||||
Db    615  ------TCAGTGAGGCAGATGACAGATATCTGTGACCGCTTCTA 659

Qy    710  CCCGATAGCCTGTGATGGTGGTTCAATTCCAGCATATAATGTGGTCTCATCCT 769
               |||||||   |||| |||||||
Db    660  CCCCAATGACTTGTGGGTGGTTGTTGTTCAGTTGCAGCACATCATGGTTGGCCTTATCCT 719

Qy    770  GCCCGGCATCGTCATCCTCCTGTTACTGACATCATCATCATCTCAAGCTGTCACACTCCAA 829
               |||||||| |||||||||||||||||||||||||||||||||
Db    720  GCCTGGTATTGTCATCCTGCTGCCATTGCATTATCATCTCCAAGCTGTCACACTCCAA 779

Qy    830  GGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGCAGTCATCCTCATCCTCAGCTTTCTTGC 889
               |||||||||||||||||||||||||||||||||||||||
Db    780  GGGCCACCAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCAGCATCCTCATCCTGGCCTTTCTTCGC 839

Qy    890  CTGCTGGCTGCCATATTATGTGGGGATCAGCATCAGCTGCTTCTTTGGGAGTCAT 949
               ||||||||||||||||||||||||||||||||||||||||||
Db    840  CTGTTGGCTGCCTACTACATTGGGATCAGCATCAGCTGCATCCTCTGGAAATCAT 899

Qy    950  CAAGCAACGAATGTGACTTCGAGAGCAATTGCACAAGTGGATCTCCATCACAGAGGCCCT 1009
               ||||||||||||||||||||||||||||||||||||||||||||
Db    900  CAAGCAACGGGTGTGAGTTTGAGCACAGTGCACAAGTGGATTTCCATCACCGAGGGCCCT 959

Qy   1010  CGCCCTTCTTCCACTGTTGCCTGCAACCCATCCTCTATGCCTTCCTCGCGGGGCCAAGTCAA 1069
               |||||||||||||||||||||||||||||||||||||||||||
Db    960  AGCTTCTTCCACTGTTGTCTGCAACCCATCCTTCTTGGAGCCAAATTTAA 1019

Qy   1070  AAGTCTGCCCAGCATGCACTCAACTCCATGAGCAGAGGCTCAACTCCAGCCTCAAGATCCTTTC 1129
               |||||||||||||||||||||||||||||||||||||||||||
Db   1020  AACTTCTGCCCAGCACCGCACTCACCTCTGAGCAGAGGGTCAGAATCCTCCAGTTTTCA 1189

Qy   1130  CAAAGGAAAGCGGGTGGCACTCTTCCGTTCCACTGAGTCTGAGTCTCAAGTTTTCA 1189
               |||||||||||||||||||||||||||||
Db   1080  CAAAGGAAAGCGAGGTGGACATTCAGCAAACCTCGAGTCTGAGTCTCAAGTTTTCA 1139

Qy   1190  CTCCAGCTAACCCTTATGCAAAGACTTATATATATATATAAGAACTTTTTTATGATAAATAACTTTT 1249
               ||     |||  |||||| ||||||||||||||||
Db   1140  CTCCAGCTAA-------CACAGATGTAAAAGACTTTTTTTATACGATAAAATAACTTTT 1191

Qy   1250  TTTTATGTTACACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTT 1309
               |||||||||||||||||||||||||||||||||||| ||
Db   1192  TTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTTTTATTGC 1251

Qy   1310  TTTTTAATTGACTGTTGGGAGTTTATGTTCCTCTAGTTTTTTGTGAGGTTTGACTTAATTT 1369
               ||                 ||||||||||||||
Db   1252  TTGTTGGATTTTGT------CTTGTGTTTTCTTAGTTTTTTAGTTTTTGTGAAGTTAAATTGACTTA 1305

Qy   1370  ATATAAATATTGTTTTTGTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGGCC 1429
               ||||| |||||||||||
Db   1306  TTTATA----TAAATTTTTTTGTTCATAATTGGATGTGTGTCTAGGCAGGACCTGTGGCC 1361

Qy   1430  AAGTTCTTAGTAGCTGTTATCGTGTGTAGGACTGTGAGACGAGAGAAACTGA 1489
               |||||||||||||||||||||||||
Db   1362  AAGTTCTTAGTTGCTGTTGTCTGTGGTAGGACTGTAGAA-------AAGGGAACTGA 1413

Qy   1490  ACATTCCAGAATGTGGTAAATTGAATAAGCCTAGCCGTTGATCCTCAGCTGTTGCTGCA 1549
               ||||||||||||||||||||||||||||||||||||||
Db   1414  ACATTCCAGAGCCGTGTAGTGACGTGAATCACGTAAAGCTAAGACTAGAAATGATCCCCAGTGTTATGCA 1473

Qy   1550  TAATCTCTCATTCCGAGGAGCACCCCACCCCACCCCCATTCTTAAAT 1609
               |||||
Db   1474  TAGATAATCTCTCCA-- 1488

Db   1610  TGTTTGGTTATGCTGTCGTGATGGTTGTTTGGTTTTTTTGTGTTGTGTTTTTTTT 1669

Qy   1670  TTTTCTGTAAAAGATGTGGCACTTAAAAACCAAAGCCTGAAATGGTGTAGAAAATGGGGT 1729
               ||||||||| |||| |||||||||  || ||||||||||  ||
Db   1526  TTTGCTGTAGAAGATGGCACTTATAACCAAAGCCCAAAGTCGGT-ATAGAAATGCTCGG--- 1581

Db   1489  ------------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGAACGTGAT 1525
```

```
RESULT 3
US-10-211-462-80
; Sequence 80, Application US/10211462
; Publication No. US2004003349SA1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-06200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-211-462-80

Query Match     50.8%;  Score 952.8;  DB 13;  Length 1679;
Best Local Similarity  80.2%;  Pred. No. 5e-209;
Matches 1210;  Conservative  0;  Mismatches 252;  Indels 46;  Gaps  6;

Qy    44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGGTCTCCGGTAACCACCACGGCTG 103
            ||  ||||  |||||||  |||  |||  ||||| |  ||| || |||||  ||||||
Db    12 GCGGCAGCAGGTAGCAAAGTGAGCCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG  71

Qy   104 TAGAGCGAGTGGTGTTGCCATGGAACCGATCAGTGTGAGTATATCGATTGATAACTACTC 163
Db    72 GAGAACCAGGGGTTACCATGGA----GGGATCAGTATATACACTTCAGATAACTACAC  126

Qy   164 TGAAGAAGTGGGGTCTGAGACTAGTTCCACCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
Db   127 CGAGGAAATGGGGCTCAGGGGACTAGTTCCACCAACAACCCCGTTTCCGTGAAGAAAA  186

Qy   224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCCACTTCATCATCTTGACTGGCAT 283
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCCACTTCTTCTAACTGGCAT  246

Qy   284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGTCATGGTCTGTCATGGTCATGACTGAGAGCATGAC 343
Db   247 TGTGGGCAATGGATTGGTCATCCTGTCATGGTCATGGTCTGTCATGGTCATGAGAAGCATGAC  306

Qy   344 GGACAAGTACCGGCTGCACCTGCACCTGACCTCCTTGTCATCACACTCCCCTT 403
Db   307 GGACAAGTACAGGGCTGCACCTGCACCTGACCTGCCGACCTCCTTGTCATCACACGCTTCCCTT  366

Qy   404 CTGGGCAGTGATGCCATGGCTGACTGGTACTTTGGTAAGGCTGTCCA 463
```

```
Qy  1730 TTTTTGTTTGTTTGTTTTTCAGTTTTCAAGAGTAGAATTGACTTCAGTCCCTACAAAT 1789
Db  1582 ------------TTTTTCAGTTTTCAGGAGTGGGTTGATTTCAGCACCTAC-AGT 1623

Qy  1790 GTACAGTCCTTGTATTACATTGTTAATAAAGTCAATGATAAAACTTAAAAAAAAAAAAAA 1849
Db  1624 GTACAGTCCTTGTATTAAGTTGTTAATAAAGTACATGTTTAAACTTAAAAAAAAAAAAAA 1683

Qy  1850 AAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
Db  1684 AAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

```
Db   367 CTGGGCAGTTGATGCCGTGGCGTGCAAACTCGGTACTTCTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCACACTGTCAACCTCTCACAGCGGCGTTCTCATCCTGGCCTTCATCCAGTCCTGA 523
Db   427 TGTCATCTCACAGTCAACCTCTCACAGCAGTCGTCCTTCATCCTGGCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTCGCCATTGTCACGGCACCACCAACAGTCAAAGGCCAAGAAACTGCTGGCTGA 583
Db   487 CCGTACCTGGCCATCGTCACGGCCACCAACAGTCAGACGCCAAGGAAGCTGTTGGCTGA 546

Qy   584 AAAGGCAGTCTATGTGGGGGTCTTGGATCCCAGCCCTCCTCCTGCTATACCTGACTTCAT 643
Db   547 AAAGGTGGTCTATGTGGGGTCTGATCCCGTGCCCTCCTGCTATTCCCGACTTCAT 606

Qy   644 CTTTGCCGACGTCAGCAGCGTCAGGGGGGACACAGTCAGGCAGGGGATGACGAGGTACATCTGTGACCG 703
Db   607 CTTTGCCACG------------TCAGTGAGGCAGATGACAGATATCTGTGACCG 651

Qy   704 CCTTTACCCGATAGCCTGTGGATGGTGGTTTCAATTCCAGCATATAATGGTGGGTCT 763
Db   652 CTTCTACCCCAATGACTGTGGGTGGTGGTGTTGTTCCAGTTTCAGCACCATCATCGTGGCCT 711

Qy   764 CATCCTGCCCGCATCGTCATCCTCTGTTACTGCATCATCATCTCTAAGCTGTCACA 823
Db   712 TATCCTGCCTGGTATTGTCATCCTGTCCTGCTATTGCATTTGCATTATCATCTCCAAGCTGTCACA 771

Qy   824 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACGCAAGGGGATGAGCAGCATTCGTGCAAGTGATCTCCATCACAGA 883
Db   772 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACCCAAGACCAGTCATCCTCATCCTCATCCTGCTTT 831

Qy   884 CTTTGCCTGCTGGCTGCCATTATATGTGGGGATCAGCATCGACATCTCCTTCATCCTTTTTGGG 943
Db   832 CTTCGCCTGTTGGCTGCCTTACTACATTGGGATCAGCATGGGCATGATCTCCATCCTCTCTGGA 891

Qy   944 AGTCATCAAGCAAGGATGTGACTTCCAGAGCATTGTGCACAAGTGATGCATCTCCATCACAGA 1003
Db   892 AATCATCAAGCAAGGGTGTGAGTTTGAGAACACACTGTGCACAGTGATTCCATCACCGA 951

Qy  1004 GGCCCTGCCTTCTTCCACTGTGCCTGCAGCATGCCTTCTCCATCCTCTATGCCTTCCTCCGGGGGCCAA 1063
Db   952 GGCCCTAGCTTTCTTCCACTGTGGCTGCAGCATGCCTTCTCCATCCTCTATGCCTTCCTCCTGGAGCCAA 1011

Qy  1064 GTTCAAAGCTCTGCCCAGCATGCACCATCACTGTCCGCTCCCGTCTTCCGCGTCCGGAGTCAGAATCCTCCAG 1123
Db  1012 ATTTAAAACCTTGCCAGCACGCGCATCACCACTCCTCCGCGTTCCACTGTTTCACTGAGTCTGAGCTGAGGT 1071

Qy  1124 CCTTTCCAAAGGAAAGCGGGTGGACACTGTGGACACTTCCGGTCCGGTGGGTGGAATCCTCCAG 1183
Db  1072 CCTCTCCAAAGAAAGCGAGGTGGACATTCATCTGTTCACTGTTTCACTGAGTCTGAGTCT 1131

Qy  1184 TTTTCACTCCAGCTAACCCTTATGCAAAGACTTAATATATATATATATATATAAA 1243
Db  1132 TTTTCACTCCAGCTAACCAGCAGATGTAAAAGACTTTTT------TTTATACGATAAATA 1183

Qy  1244 GAACTTTTTATGTTACACATTTCCAGATAAGAGACTGACCAGTCTTGTACAGTTTT 1303
Db  1184 ACTTTTTTTTAAGTTACACATTTTTCAGATAAAAGACTGACCAATATTGTACAGTTTT 1243

Qy  1304 TTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCCTGTAGGTTTGGACT 1363
Db  1244 TATTGCTTGTTGTTGATTTTGT-----CTTGTGTTTCTTTAGTTTTGTTGTTTAGAAGTTTAATT 1297

Qy  1364 TAATTATATAAAATATGTTTTTTGTTTTCATGTGAATGAGCGTCTAGGCAGGACCT 1423
Db  1298 GACTTATTTATA---TAAATTTTTTTGTTTTCATTGATGTGTTCTAGGCAGGACCT 1353

Qy  1424 GTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGGTGGACTGTAGAACTGTAGAGAGAAGA 1483
Db  1354 GTGGCCAAGTTCTTAGTTGTGCTGTATGGTCTCGTGGTGTGGTGGGACTGTGTGAGA-------AAGGG 1405

Qy  1484 AACTGAACATTCCAGAAGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTT 1543
```

Db 1406 AACTGAACATTCCAGAGCCGTGTAGTGAATCACGTAAAGCTAGAAATGATCCCCAGCTGTT 1465

Qy 1544 GCTGCATA 1551
              ||||||||
Db 1466 TATGCATA 1473

RESULT 4
US-10-181-906-9
; Sequence 9, Application US/10181906
; Publication No. US20040053864A1
; GENERAL INFORMATION:
; APPLICANT: Karsenty, Gerard
; APPLICANT: Amling, Michael
; APPLICANT: Ducy, Patricia
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BONE FORMATION VIA
; TITLE OF INVENTION: MODULATION OF NEUROPEPTIDE Y ACTIVITY
; FILE REFERENCE: 9142-020-999
; CURRENT APPLICATION NUMBER: US/10/181,906
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/02040
; PRIOR FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 9
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-181-906-9

Query Match   50.8%; Score 952.8; DB 13; Length 1679;
Best Local Similarity   80.2%; Pred. No. 5e-209;
Matches 1210; Conservative   0; Mismatches 252; Indels 46; Gaps   6;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
            ||||||||||||||||||| |||||||||||||| ||||||||||||||| ||||||
Db   12 GCGGCAGCAGGTAGCAGAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCCATCTG   71

Qy 104 TAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
            || |||||||||| ||||||||||| |||||||| |||||||||||||||||||||||
Db   72 GAGAACCAGCCGTTACCATGGA-----GGGGATCAGTAGTATACACTTCAGATAACTACAC 126

Qy 164 TGAAGAAGTGGGGTCTTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
            ||||||||||||| ||||||||||| |||||||| ||||||||||||| ||||||| |
Db 127 CGAGGAAATGGGCTCAGGGGACCTGCAATGACTCCATGAGGAACCCTGTTTCCGTGAAGAAA 186

Qy 224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCACTCTTCTTGACTGGCAT 283
            |||||| |||||||| |||||||||| |||||||||||||||||| ||||||||||||
Db 187 TGCTAATTTCAATAAAAATCTTCCTGCCCACCATCTACTCCATCACTTCTTAACTGGCAT 246

Qy 284 AGTCGGCAATGATTGTGATCCTGGTCATGGTTACCAGAGAGAGCTAAGGAGCATGAC 343
            |||||||||||||||| ||||||||||||||||||||||||||| |||||||||||||
Db 247 TGTGGGCAATGATTGGTCATCCTGGTCATGGTTACCAGAGAACTGAGAAGCATGAC 306

Qy 344 GGACAAGTACCGGCTGCACCTGTGCACCTGGTCGACCTCCTCTTGTCATCACACTCCCCTT 403
            ||||||||||||||||||| ||||||| | ||||||||||| ||||||||||||||||||
Db 307 GGACAAGTACAGCGCTGCACCTGCACCTGTCAGTGGCCGACCCTCCTCTTGTCATCACGCTTCCCTT 366

Qy 404 CTTGGGCAGTTGATGCCATGGTACTTTGGGAAATTTTTGTGTTGTAAGGCTGTCCA 463
            ||||||||||||||||||| ||| |||| |||||| ||||||||||||||||||||
Db 367 CTTGGGCAGTTGATGCCGTGGCAAACTTCTTACACAGCAGTGTCCTCATGCAGAGGAGTCCA 426

Qy 464 TATCATCTACACTGTCAACCTGTCCCGTT 523
            ||||||||||| ||||||||||||||||||
Db 427 TGTCATCATCACACAGTCAACCTTCATCAGTCTGGA 486

Qy 524 CCGGTACCTCGCCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGAAAACTGCTGCTGA 583
            ||||||||||||||||| |||||||||| ||||||||||||||||||||||||||||||
Db 487 CCCCTTACCTGGCCATCGTCCACCGCCACCAAACAGTCAGAGGGCCAAGGAAGCTGTTGGCTGA 546

Qy 584 AAAGGCAGTCTATGTGGGCGTCGGAGTCCCCAGCCCTCCTGACTATACCTGACTTCAT 643
            |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db 547 AAAGGTGGTCATAGTTGGCGTTCGGATCCCCTGCCCCTGCTGACTATTCCCGACTTCAT 606

Qy 644 CTTTGCCGACGTCAGCCAGGAGATCAGTCAGGGGGATCAGTCAGGGGATGACAGGTACATCGTGACCG 703
            |||||||||||||| ||||||||||-----------TCAGTGAGCGCAGATATCTGTGACCG
Db 607 CTTTGCCAACG-----------TCAGTGAGCGCAGATATATCTGTGACCG 651

Qy 704 CCTTACCCGATAGCCTGTGGATGGTGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCT 763
            || ||||||||||||||||||||||| ||||||||||||||||||||| ||||| ||||
Db 652 CTTCTACCCCAATGACTTGTGGGGTGGTGTTGTTCCAGTTTCAGTTCAGCACATCATGGTTGGCCT 711

Qy 764 CATCCTGCCGGCATCGTCATCCTCCTGTTACTGCATCATCATCTAAGCTGTCACA 823
            |||||||||||| |||||||||||||||| |||||||||||||||||||||||||
Db 712 TATCCTGCCTGGTATTGTCATCCTGTCCTGCTATTGCATTATTGCATCATCCTCCAAGCTGTCACA 771

Qy 824 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTT 883
            ||||||||| ||||||||||||||||||| |||||||||||||||||||||||||||
Db 772 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCATCCTCTGGCTTT 831

Qy 884 CTTTGCCTGCTGGCTGCCATATATGTGGGGATCGGACATCGACTCCTTCATCCTTTGG 943
            ||||| ||| ||||||||| |||||||||| ||||||| ||||||| |||||||||||
Db 832 CTTCGCTGTTGGCTGCCTTACTACATTGGGATCGATCAGCATCGACTCCTTCATCCTCCTGA 891

Qy 944 AGTCATCAAGCAAGGATGTGACTTCGAGACATTGTGCACAAGTGGATCTCCATCACAGA 1003
            ||||||||| ||||||||| ||||||||||||||| |||| |||||||||||||||||
Db 892 AATCATCAAGCAAGGGGTGTGAGTTTGAGTTTGAGAACAACACTGTGCACAAGTGGATTCCATCACCGA 951

Qy 1004 GGCCCTGCCTTCTTCCACTGTTGCCTGAACCCATCCTCATGCCCTCCTGCCGGGGCCAA 1063
            ||||||||||||||| ||||||||||||||| |||||||||||||||||| |||| |||
Db 952 GGCCCTAGCTTCTTCCACTGTTGCTGAACCCCATCCTCCTATGCCTCTGGGAGCCAA 1011

Qy 1064 GTTCAAAAGCTCTGCCCAGCATGCACCATCCAACTCACTGCAGAGAGGCTCCAGCCTCAAGAT 1123
            |||||||||||||||||||||| |||||| |||||||| |||||||||||||||||||||||
Db 1012 ATTTAAAACCTCTGCCCAGCACGCACCACTCACCCTCTGTGAGCAGAGGGTCCAGCTTCAAGAT 1071

Qy 1124 CCCTTCCAAAGGAGAAGCCGGGGTGACACTCTTCCGTCTCCACGGAGTCGAAATCCTCCAG 1183
            |||||||||||||||||| |||||||||||| || ||||| |||||||||| | |||||
Db 1072 CCTCTCCAAAGGAGAAGCAGGGTGGCACATTTCATCATCTGTTTCCACTGAGTCTTCAAG 1131

Qy 1184 TTTTCACTCCAGTAACCTTATGCAAAGACTTATATATAATATATATAGATAAA 1243
            ||||||||||||||| |||||||||||| ||||||| |||||||||
Db 1132 TTTTCACTCCAGTAACACAGATGTAAAGAGACCTTTTT-------TTTTACGATAAAATA 1183

Qy 1244 GAACTTTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT 1303
            |||||||||| |||||||||||||| ||||||| ||||||||||| |||||| |||||||
Db 1184 ACTTTTTTTTAAGTTACACATTTTTCAGATATAAAGAGACTGACCAATATTGTACAGTTTT 1243

Qy 1304 TTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCCCTCTAGTTTTGTGAGGTTTGACT 1363
            || ||||||||||||| ||||||||||||||||||| ||||||| |||||||||||
Db 1244 TATTGCTTGTTGAGATTTTGT-----CTTGTGTTTCTTTAGTTTTGTTTGTGAAGTTTAATT 1297

Qy 1364 TAATTTATATAAATATTGTTTTTTTGTTTGTTCATGTGAATGAGGGTCTAGGCAGGACCT 1423
            |||||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||
Db 1298 GACTTATTTATA----TAAATTTTTTTGTTTCATATTGTGTCATATTGTGATGGACTGTGAGAAGTGTGTCTAGGCAGGACCT 1353

Qy 1424 GTGGCCAAGTTCTTAGTAGCTGTTTATCGTGTGTGTAGAACTGTGAGAGGGAAGA 1483
            |||||||||||||||||||||||||| || || ||||||||||||||||| -------AAGGG
Db 1354 GTGGGCCAAGTTCTTAGTTGCTGTATGTCTTCGTGTGGGTAGGACTGTAGAAGA 1405

Qy 1484 AACTGAACATTCCAGAGCTGTGTAGTGAATGAATTGAATGAAATAAAGCTAGAAATGATCCCCAGCTGTT 1543
            |||||||||||||||||| |||||| ||||| ||||||||||||||||||||||||||||||||||||
Db 1406 AACTGAACATTCCAGAGCCGTGTAGTGAATCACGTAGTGAATCACGTAAAGCTAGAAATGATCCCCAGCTGTT 1465

Qy 1544 GCTGCATA 1551
            ||||||||
Db 1466 TATGCATA 1473

RESULT 5
US-10-342-887-912
; Sequence 912, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong

```
;   APPLICANT:  Linsley, Peter S.
;   APPLICANT:  Mao, Mao
;   APPLICANT:  Roberts, Christopher J.
;   APPLICANT:  Van 't Veer, Laura Johanna
;   APPLICANT:  Van de Vijver, Marc J.
;   APPLICANT:  Bernards, Rene
;   TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;   FILE REFERENCE: 9301-188-999
;   CURRENT APPLICATION NUMBER: US/10/342,887
;   CURRENT FILING DATE: 2003-01-15
;   PRIOR APPLICATION NUMBER: 60/299,918
;   PRIOR FILING DATE: 2001-06-18
;   PRIOR APPLICATION NUMBER: 60/380,710
;   PRIOR FILING DATE: 2002-05-14
;   PRIOR APPLICATION NUMBER: 10/172,118
;   PRIOR FILING DATE: 2002-06-14
;   NUMBER OF SEQ ID NOS: 2699
;   SEQ ID NO 912
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-912

Query Match        50.8%;   Score 952.8;   DB 13;   Length 1679;
Best Local Similarity   80.2%;   Pred. No. 5e-209;
Matches 1210;   Conservative   0;   Mismatches 252;   Indels  46;   Gaps   6;

Qy    44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCGTTTGGTGCTCCGGTACCACCACGGCTG 103
Db    12 GCGGCAGCAGGTAGCAAAGTGACGGCCGAGGGCCTCGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy   104 TAGAGCGAGCAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
Db    72 GAGAACCAGCCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACTAC 126

Qy   164 TGAAGAAGTGGGGTGGACAGCATGACTCCAACAAGAACCCTGCTTCCGGGATGAAAA 223
Db   127 CGAGGGAAATGGGGCTCAGGGGACTAATGACTCATGAAGGAACCCCGTTTCCGTGAAGAAAA 186

Qy   224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTGACTGGCAT 283
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTCCATCATCTTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAGAGAGCTAAGGAGCATGAC 343
Db   247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAGAGAAGAAACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCTCTTTGTCATCACACTCCCCTT 403
Db   307 GGACAAGTACAGCGCTGCACCTGTCAGTGGCCGACCTCCTTTGTCATCACGCTTCCCTT 366

Qy   404 CTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
Db   367 CTGGGCAGTTGATGCCGTGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTCTACAGCGCCACCACAACAGTCAAAGGCCAAGGAAACTGCTGGA 523
Db   427 TGTCATCTACACAGTCAACCTCTACAGCACAGTGTCCCTTCATCCTGGCCTTCATCAGTCGTCTGGA 486

Qy   524 CCGGTACCTGCCAATTGCCACAAGTGCCAAGCCGGTCTGATCCCAGAGAAACTGCTGGCTGA 583
Db   487 CCGCTACCTGCCATCGTCCACACGTCCACACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy   584 AAAGGCAGTCTATGTGGGCGTCTGGATCCCCAGCCCTCCTCCTGACTATACCTGACTTCAT 643
Db   547 AAAGGTGGTCATGTTGGGCGTCTGGATCCCGCCCTCCTGCTGACTAATTCCCGACTTCAT 606

Qy   644 CTTTGCCGACCTCAGCCCAGCCAGGGGGACAATCAGTCAGGGGGATGACAGGGTACATCGTGACCG 703
Db   607 CTTTGCCAACG---------TCAGTGAGGCAGATGACAGATATCGTGACCG 651

Qy   704 CCTTTACCCCGATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAAATGGTGGGTCT 763
Db   652 CTTCTACCCCAATGACTTGTGTGGGTCGGTGTTCCAGTTCAGCGACACATCATGTGGTTGGCCT 711

Qy   764 CATCCTGCCCGGCATCGTCATCCTCCTGTTACTGCATCATCATCTTAAGCTGTCACA 823
Db   712 TATCCTGCCTGTATTGTCATCCTGCTGTCTATTGCATTATCATCTCCAAGCTGTCACA 771

Qy   824 CTCCAAGGGCACCACAGAAGGCAAGGCCCTCAAGACGCAGCAGTCATCCTCATCCTCAGCTTT 883
Db   772 CTCCAAGGGCACCACAGAAGGCAAGGCCCTCAAGACCACAGTCATCCCTCATCCTGGCTTT 831

Qy   884 CTTTGCTGCTGCTGGCTGCCATATATGTGGGGATCAGCATCGACTCCTTCATCCTTTTGG 943
Db   832 CTTCGGCCTGTTGGCTGCCTACTACATTGGATGGATCAGCATCGACTCCTTTCATCCTCCTGA 891

Qy   944 AGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACACGA 1003
Db   892 AATCATCAAGCAAGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTCCATCACCGA 951

Qy  1004 GGCCCTCGCCTTCTTCTCCACTGTGCCTGAACCCCATCCTCTATGCCTTCCTGCGGGGCCAA 1063
Db   952 GGCCCTAGCTTCTTCTCCACTGTGCTCTGAACCCCCATCCCTCTATGCTTCCTTGGAGCCAA 1011

Qy  1064 GTTCAAAAGCTCTGCCCAGCATGACCACTCCATGGAGCAGAGGCTCAAGAT 1123
Db  1012 ATTTAAAACCTCGCCCAGCAGCGGACACTGCTCCTCGTGAGCAGAGGGTCCAGCCTCAAGAT 1071

Qy  1124 CCTTTCCAAAGGAAAGCGGGTGGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAG 1183
Db  1072 CCTCTCCAAAGGAAAGCGAGTGGACATTCATCGTTCCACTGAGTCTGAGTCTTCAAG 1131

Qy  1184 TTTTCACTCCAGCTAACCCTTATGCAAAGACTTATATATATATATATATGATAAA 1243
Db  1132 TTTTCACTCCAGCTAACACAGATGTAAAAGACTTTTT------TTTATACGATAAATA 1183

Qy  1244 GAACTTTTTATGTTACACATTTCCAGATATAGAGACTGACCAGTCTCTGTACAGTTTT 1303
Db  1184 ACTTTTTTTAAGTTACACATTTTTCAGATATAAAAAGACTGACCAATATTGTACAGTTTT 1243

Qy  1304 TTTTTTTTTTTAAATTGACTGTTGGGAGTTTATGTTCCTCCTAGTTTTTGTGAGGTTTGACT 1363
Db  1244 TATTGCTTGTTGGGATTTTTGT------CTTGTGTTTCTTTAGTTTTGTGAAGTTAATT 1297

Qy  1364 TAATTTATATAAAATATTGTTTTTTTTGTTCATGTGAATGAGCGCTCAGGCAGGACCT 1423
Db  1298 GACTTATTTATA----TAAATTTTTTTTTGTTTCATATTGATGTGTGTCTAGGCAGGGACCT 1353

Qy  1424 GTGGCCAAGTTCTTAGTAGCTGTTTATCGTGTGTAGGACTGTAGAACTGTAGAGGAAGA 1483
Db  1354 GTGGCCAAGTTCTTAGTTGCTGTGTATGTCTCGTGTGGTAGGACTGTAGAA-------AGGGG 1405

Qy  1484 AACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCCTCAGCTGTT 1543
Db  1406 AACTGAACATTCCAGAGCGTGTAGTGAATCACGTAAAATGATCCCCAGCTGTT 1465

Qy  1544 GCTGCATA 1551
Db  1466 TATGCATA 1473
```

```
RESULT 6
US-10-151-274-14
;   Sequence 14, Application US/10151274
;   Publication No. US20030064071A1
;   GENERAL INFORMATION:
;   APPLICANT:  Littman, Dan R.
;   APPLICANT:  Kwon, Douglas S.
;   APPLICANT:  van Kooyk, Yvette
;   APPLICANT:  Geijtenbeck, Theo
;   TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
;   TITLE OF INVENTION: INTO
;   TITLE OF INVENTION: CELLS
;   FILE REFERENCE: 1049-1-017
;   CURRENT APPLICATION NUMBER: US/10/151,274
```

```
;  CURRENT FILING DATE: 2002-05-20
;  PRIOR APPLICATION NUMBER: US/09/517,605
;  PRIOR FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 14
;  LENGTH: 1679
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-151-274-14

Query Match         50.8%;  Score 952.8;  DB 13;  Length 1679;
Best Local Similarity  80.2%;  Pred. No. 5e-209;
Matches 1210;  Conservative 0;  Mismatches 252;  Indels 46;  Gaps  6;
```

```
Qy       44  GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTG  103
              || ||||||||||||||||| |||||  |||| |||||||| |||||||||||||||||
Db       12  GCGGCAGCAGGTAGCAAAGTGACGGCCGAGGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG   71

Qy      104  TAGAGCGAGTGTTGCCATGAACCATCAGTGTGAGTATATACACTTCTGATAACTACTC  163
              |||||  | || |||  ||||  ||| | |||||  ||||| ||  ||||||||||||
Db       72  GAGACCAGCGGGTTACCATGA------GGGGATCAGTAGTATACACTTCAGATAACTACAC  126

Qy      164  TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA  223
              |  | ||||  | |||||||||||||||||  ||||||||||  | ||  ||||| ||||
Db      127  CGAGGAAATGGCTCAGGGGACTATGACTCCATGAGAGGAACCCTCGTTTCCGTGAAGAAAA  186

Qy      224  CGTCCATTCAATAGGATCTTCCTGCCCACCATCACTTCACTTCCTCCTGACTGGCAT  283
              |||||| || || |||||||| |||||| ||||| |||||||| || ||||||||||
Db      187  TGCTAATTTCAATAAAATCTTCCTGCCCACCATTCACTTCCATCATCCTCCTCAACTGGCAT  246

Qy      284  AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGGAAGAGCTAAGGAGCATGAC  343
              ||||||| ||||| ||| |||||||||||| ||||||||||||| ||||| ||||||||
Db      247  TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGGAAGAACTGAAGCATGAC  306

Qy      344  GGACAAGTACCGGCTGCACCTGTCAGTGGGTGACCTCCTCTTGTCATCACACTCCCTT  403
              || |||||| ||||||||||||| |||||||||||||| | ||| | |||||||| ||
Db      307  GGACAAGTACAGGCTGCACCTGTCACTGGGGTACTTTGGGAAATTTTGTGTAAGGCGTCCA  366

Qy      404  CTGGGAGCAGTTGATGCCATGGCCGTACTTGGAAACTCCATGCCTTCATCCTGGCCTTCA  463
              |||||||||||||| |||||| |||||||||||||||| ||| || ||||| |||||||
Db      367  CTGGGCAGTTGATGCCTGGCCGTACTTGGAAACTCCATGCCTTCATCAGTCGTCTGGA  426

Qy      464  TATCATTCTACACTGTCAACCTCTACAGCAGCGTTCCCATCCTGGCCTTCATCAGCCTGGA  523
              ||||| |||| |||| | |||| ||||| |||||| ||  |||||||||||||||||||
Db      427  TGTCATCTACACACAGTGACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA  486

Qy      524  CCGGTACCTCGCCATTGTCCACGCCCAACAGTCCAACAGTCAAAGGCCAAGGAAACTGCTCGGCTGA  583
              ||||||||||  |||||||||| ||| ||||||||||||||||||||||||||||||| || ||||
Db      487  CCGCTACCTGGCCATCGTCCACGCCGTCGGATCCCAGCCCTCTGACTTCCGACTTGACTGTCGGCTGA  546

Qy      584  AAAGGCAGTCTATGTGGGCGTCTGGATCCCTGCTGCTGCTCCTGACTGACTTCCCGACTTCAT  643
              |||||| || |||||||||||||||||||||| | ||||||||||||||||||||||||||
Db      547  AAAGGTGCTCATGTGGCGTCTGGACTATTGTGGCTTGTCATCCTGTTGCATTATCCCGACTTCA  606

Qy      644  CTTTGCCGACGTCAGCAGGGGGACATCAGTCAGGGGATGACAGGTACATCGTGACCG  703
              || |||||| || ||    ||||||||||||||||||||||||||||||||||||||
Db      607  CTTTGCCAAGG-----TCAGTGAGGCAGATGACAGATATCGTGACCG  651

Qy      704  CCTTTACCCGGATAGCGTCTGGATGCGTGGTGTTCAATTCAATTCCAGCATATAATGTGGGTCT  763
              || | |||||| |||||  ||||| |  || ||||||||| ||||  ||| || ||||||||||
Db      652  CTTCTACCCCAATGACTTGTGGGGTGTGTGTCCAGTTTCAGCAGCACATGATTGTTGGCCT  711

Qy      764  CATCCTGCCCGGGCCATCGCGTCATCCTGTTACTGCAATCATCATCCTAAGGCTGTCCACA  823
              |||| | | ||||| | ||||| ||||||||| || | |||||||||||||||||||||||
Db      712  TATCCTGCTGGTATTGTCATCCTCGTATTATATTCATTCCTCCCAAGCTGTCTGGCTTTCACA  771

Qy      824  CTCCAAGGGCCACCAGAAGCCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTGCTAGCTTT  883
              || |||  ||||||||||| |||||||| ||||||| ||||  ||| |||||| || |||||||
Db      772  CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACCACCACAGTCATCCTCATCCTGGCTTT  831

Qy      884  CTTTGCCTGCTGGTGGGTGCCGCCATATATGTGGGATCGAGCATGCATGGTGTTTGGG  943
              |||| |||||| |||| || ||| ||| |||||| ||||||| ||||| |||||||||||
```

```
Db      832  CTTCGGCCTGTTGGCTGGCTGCCTTACTACATTGGATCAGCATCGACTCCTTCATCCTCCTGGA  891

Qy      944  AGTCATCAAGACAAGGATGTGACTTCGAGAGACCATTGTGCACAAGTGATCTCCATCACAGA  1003
              | |||  |||| |||||  ||  || || |||||||||||||| ||||| |||||||||||
Db      892  AATCATCAAGACAAGGGTGTGAGTTTGAGAACACTGTCCACAAGTGGATTCCATCACCCGA  951

Qy     1004  GGCCCTGCCTTCTTCCACTGTTGCCTGGTGCGAACCCATCCTCATGCCTTCCTCGGGGGCCAA  1063
              |||||| |||||||||||| ||||| | || || ||||||| |||  || ||||||||||||
Db      952  GGCCCTAGCTTCTTCCACTGTTCTTGCTGGAACCCCATCCTCATGTGTCCTTCCTTGGGAGCCAA  1011

Qy     1064  GTTCAAAAGCTGTGCCAGCATGCACTCAACTCAACTCCATGAGCAGGAGGCTCCAGCCTCAAGAT  1123
              |||| || | |||| || || |||||| |||||||| |||||||||| |  ||||||||||||
Db     1012  ATTTAAAACCTCGCCCAGCACCCACCTCACCCTGTGGAGCAGGGGTCCAGCCTCAAGAT  1071

Qy     1124  CCTTTCAAAGGAAAGGCGGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAG  1183
              |||||||| | | |||| |  |||||| ||  ||| | | |||||||| |||||||||||||
Db     1072  CCTCTCCAAAGAAGGAAAGGGAGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAG  1131

Qy     1184  TTTTCACTCCAGTCAACCCTTATGCAAAGACTTCAAAGACTTATATATATATATGATAAA  1243
              ||||||||||||||  |||||| | ||| || |        ||| |   ||| ||||||
Db     1132  TTTTCACTCCAGTTAACACAGATGTAAAAGACTTTT------TTTATACCGATAAATA  1183

Qy     1244  GAACTTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT  1303
              ||||||| ||  || |||||||| |||| ||| |||| |||| |||| |||||||||||
Db     1184  ACTTTTTTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAAATATTGTACAGTTTT  1243

Qy     1304  TTTTTTTTTTTAATTGACTGTTGTGGGAGTTTATGTTCCCTAGTTTTTGTGAGGTTTGACT  1363
              |||||||||||   || | || ||        ||||||      | ||| || |||||||
Db     1244  TATTGCTTGTTGGATTTTTGT------CTTGTGTTTTCTTTAGTTTTTTGTGAAGTTTAATT  1297

Qy     1364  TAATTTATATAAAATTGTTTTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCGAGGACCT  1423
              |||| || || | ||| ||||||  | || |||| || | |   | || |||| ||||||
Db     1298  GACTTATTTATA---TAAATTTTTTTGTTTCATATTGATGTGTGTCTAGGCAGGACCT  1353

Qy     1424  GTGGCCAAGTTCTTAGTAGCTGTTATCGTGTGTAGGAACTGTAGAGGAAGA  1483
              || ||||||||||||| || ||||||| ||| |   | | |   | ||||
Db     1354  GTGGCCAAGTTCTTAGTTGCTGCTGTATGTCTGTATGCTCGTGGTGGAGAA-------AAGGG  1405

Qy     1484  AACTGAAACATTCCAGAATGTGTGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTT  1543
              |||||| |||||||||||  |||||||||| ||||||||| ||||||||  |||||||||
Db     1406  AACTGAACATTCCAGAGCCGTGTAGTGAATCACGTAGAAATGATCCCCGCTGTT  1465

Qy     1544  GCTGCATA  1551
              ||||||||
Db     1466  TATGCATA  1473
```

```
RESULT 7
US-10-172-118-912
; Sequence 912, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT: He, Yudong
;  APPLICANT: Linsley, Peter
;  APPLICANT: Mao, Mao
;  APPLICANT: Roberts, Chris
;  APPLICANT: van 't Veer, Laura
;  APPLICANT: Van de Vijver, Marc
;  APPLICANT: Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-175-999
;  CURRENT APPLICATION NUMBER: US/10/172,118
;  CURRENT FILING DATE: 2002-06-14
;  PRIOR APPLICATION NUMBER: 60/380,770
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 2699
;  SEQ ID NO 912
;  LENGTH: 1679
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;  DATABASE ACCESSION NUMBER: NM_003467
```

```
;  DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-912

Query Match      50.8%;  Score 952.8;  DB 13;  Length 1679;
Best Local Similarity  80.2%;  Pred. No. 5e-209;
Matches 1210;  Conservative  0;  Mismatches 252;  Indels  46;  Gaps   6;

Qy    44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTCGGTGTCCGGTAACCACCACGGCTG 103
         |||| || ||||||||| | ||||  | |||||  ||||| || |||| ||| || |||
Db    12 GGGGCAGCAGGTAGCAGAAGTGACGCCGAGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy   104 TAGAGGCAGTGTGTCCATGGAACCGGATCCGACCAGTGTGAGTATATACACTTCTGATAACTACTC 163
         | |||| || |||| |||||    |   ||| || ||| || |||||| ||| ||| |||| ||| ||| ||
Db    72 GAGAACCAGCAGCGGTTACCATGGA---GGGGATCAGTAGTATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGGAGACTCGGCGGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAGA 223
         ||||| || ||| || ||| || |  | |||  |||  ||| |||||||| ||||| ||||| || |||
Db   127 CGAGGAAATGGGCTCAGGGGACCATGAGCTCCATGAAGGAACCCTGTTTCCGTGAAGAAA 186

Qy   224 CGTCCATTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCCTCTGACTGGCAT 283
         ||| |||||||||| |||||||| |||||||| ||| ||||||||| || |||||||||
Db   187 TGCTAATTTCAATAAAATCTTCCCGCCACCAACTCCATCATCTTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC 343
         | | |||||||||| || |||||| || || || |||||| || | || ||||| ||||||||
Db   247 TGTGGGCAATGGATTCATCCTGGTCATCCTGGTTACCAGAGAAACTGAGAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCGACCTGTCAGTGGCTGGCTGACCTCCTCTTGTCATCACACTCCCCTT 403
         |||||||||| || |||||||||||||||| |  |||||||||||||||||||| | ||| |||||
Db   307 GGACAAGTACAGGCTGCGACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCTTCCCTT 366

Qy   404 CTGGGCAGTGATGCCTGACCTGGCTGCTACTGTGTAAGGCTGTCCA 463
         ||||||||| ||||||| || ||| ||| ||||| || ||||| |||| |||
Db   367 CTGGGCAGTTGATGCCGTGCGTGCAAACTGGTACTTTGGGAAACTTCCCATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACGTGCAACCTCTACAGCAGCGGTTCTCATCCTGGCCCTTCATCAGCCTGGA 523
         | ||||||||||  | || ||||||||||| |  |  ||| ||| ||| ||||||||| |||||
Db   427 TGTCATCTACACAGCACCTCTACAGCAGTGTCCTCACGTGTCCCTCATCCTGGCCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTCGCTCGCCATTGCCATGTGGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTTCAT 583
         ||| |||||  || |||||||||||||||||||||||||||| |  | |||||||||||||| |||||| |||
Db   487 CCGCTACCTGGCCATGGCCATGTGGGGCGTCTGGATCCCACCGCCGCTCCTCCTGACTATACCTTCAT 546

Qy   584 AAAGGCAGTCTATGTGGGGCGTCTCGGATCCCTGCTATTCCCGACTTCAT 643
         ||||  ||||||||| |||||||| |||||||||||| || || |||||||
Db   547 AAAGGTGGTCTATGTTGGGCGTCTCGGATCCCTGCTGTTCCGGACTTCAT 606

Qy   644 CTTTGCCGACGTCAGCGCAGGGGGACATCAGTCAGGGATGACAGGTACATCTGTGACCG 703
         |||||||  ||||         |||||  |||||| || || | | | |||| ||||||
Db   607 CTTTGCCAACG----------TCAGTGAGGCAGAGTGACAGATATCTGTGACCG 651

Qy   704 CCTTTACCCCGATAGCCGTGGGATTTGTTCATTCCAGCATATAATGGTGGGTCT 763
         | ||| ||||| || | ||||| |||| ||| | ||||| |  ||||| |||| ||
Db   652 CTTCTACCCCCAATGACTTGTGGGTTGGTTGTTCAGCAGCACATCATGGTTGGCCT 711

Qy   764 CATCCTGCCCGCCGGCATCGTCATCTCTCGTACTGCATCATCATCCTAAGCTGTCACA 823
         |||||||| | | ||  ||||||||||| ||| |||| ||||||| | ||||||||||
Db   712 TATCCTGCTGGTGGTATTGTCATCTGCTATTGCATTCCAAGCTGTCACA 771

Qy   824 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACGACAGTCATCTCATCCCTAGCTTT 883
         |||||||||||||||||||||||||||||||||||  |  ||||||| | || |||| |||
Db   772 CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACCATCATCCTCCTGGCTTT 831

Qy   884 CTTTGCTGCTGGCTGCCATATTATGTGGGGATCAGCATCGACTCCTTTTGGG 943
         ||| || | ||||||| | | | | ||| |||||||||  | ||||| |||
Db   832 CTTCGCCGTGTTGGCTGCCTTACTACAATTGGGATCAGCATCATGGGATCAGCATCCTCCTGGA 891

Qy   944 AGTCATCATCAAGCAAGGATGGATGTGACTTGACTTGCCTGAGAGCATTGTGCACAGTGGATTTCCATCACCGA 1003
         | |||||| |||| ||||  ||||    ||  |||  || | || |||| ||| | |||| || ||||||||||||
Db   892 AATCATCAAGCAAGGTGTGAGTTTGAGAACACTGTGCCTCAGTGGATTTCCATCACCGA 951

Qy  1004 GGCCCTAGCTTCTTCCACTGTTGTGCTGACCCCCATCCTCTATGCCTTCCTCTCGGGGCCAA 1063
```

```
Db   952 GGCCCTAGCTTTCTTCCACTGTTGTCTGAACCCCATCCTCTATGCTTTCCTCTGGAGCCAA 1011

Qy  1064 GTTCAAAAGTCTGCCCAGCATGCGACTCCATGAGCAGGCTCCAGCCTCAAGAT 1123
         ||| |||| | ||||||||||  |||||| ||||| |||| ||| |||||||||||||||||||
Db  1012 ATTTAAAACCTCTGCCCAGCACGGCACTCACCTGTGTGAGCAGAGGGTCCAGCCTCAAGAT 1071

Qy  1124 CCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAG 1183
         ||| ||||||||||||||| | ||||| || | ||| |||||| |||||||||| | | |||
Db  1072 CCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTTCTTCAG 1131

Qy  1184 TTTTCACTCCAGCTAACACCCTTATGCAAAGACTTATATATATATATATATGATAAA 1243
         ||||||||||||||||||| ||| | ||||||||||||      ||| || ||||||  |
Db  1132 TTTTCACTCCAGCTAACACAGATGTAAAAGACTTTTT-----TTTATACGATAAATA 1183

Qy  1244 GAACTTTTTATGTTTACACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT 1303
         |||||||||| ||| |||||||| |||||||||| |||||||||| | ||||||||||||
Db  1184 ACTTTTTTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTT 1243

Qy  1304 TTTTTTTTTTTAATTGACTGTTGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTGACT 1363
         |   |||| || || ||| ||||||| |||| | |  ||||||||||||||| || | | |
Db  1244 TATTGCTTGTTGGATTTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATT 1297

Qy  1364 TAATTATATAAATATTGTTTTTGTTTGTTCATGTGAATGAGCGTCTAGGCAGGACCT 1423
         |||||||||| | |   | |||||||||||||| | ||||| || | ||||||||||||
Db  1298 GACTTATTTATA----TAAATTTTTTTGTTTTCATTATTGATGTGTGTCTAGGCAGGACCT 1353

Qy  1424 GTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTTAGGACTGTGAGAACTGTAGAGGAAGA 1483
         |||||||||||||||||| ||| | |||| | ||| |||||||||      | ||| ||
Db  1354 GTGGCCAAGTTCTTAGTTGCTTAGTTGCTGTATGTCTCGTGGTAGGACTGTGAA------AGGG 1405

Qy  1484 AACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTT 1543
         |||||||||||||||||  |||| || |||  || ||||||| | ||||||| ||||||||
Db  1406 AACTGAACATTCCAGAGCGTGTAGTGAATCACGTAAAGCTAGAAATGATCCCCAGCTGTT 1465

Qy  1544 GCTGCATA 1551
         | ||||||
Db  1466 TATGCATA 1473
```

```
Qy    44  GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG  103
          ||||| |||||||||||||||  | ||||  ||||  ||||||||||| |||||| ||||
Db    12  GCGGCAGCAGCAGTAGCAAAGTGACCGCGAGGGCCTGAGTGCTCCAGTAGCACCGCATCTG  71

Qy   104  TAGAGGCGGAGTGTTGCCATGA----GGGGATCAGTACACATTTCTGATAACTACTC  163
          |||||  | | || ||||||||     ||||||||||| |||||||||| ||||||||
Db    72  GAGAACCCAGCAGCTGGTTACCATGA----GGGGATCAGTATATACACTTCAGATAACTACAC  126

Qy   164  TGAAGAAGTCGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAA  223
          |||||||  ||||||||||  |  ||| ||||||||| |||||||||||| ||||||||
Db   127  CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGAACCCATGACCAATATTTGTGAAGAAA  186

Qy   224  CGTCCATTCAATAGGATCTTCCTGCCACCATCACTTCATCATCTTCTTGACTGGCAT  283
          ||||||||| ||| ||||| ||||| |||||||| ||||||||||| ||| |||||||
Db   187  TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTTCCTCCACCATCCTTCTTAACTGGCAT  246

Qy   284  AGTGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC  343
          |||||||||||| | |||||||| ||||||||| |||| ||||||||||||||||||||
Db   247  TGTGGGCAATGGATTGGTCATCCTGTCATGGGTTACCAGGAGAAGCTGAGAAGCATGAC  306

Qy   344  GGACAAGTACCGGCGCTGCACCTGCTCAGTGGCTGACCTGCCTCCTTGTCATCACACTCCCCTT  403
          ||| |||||||  ||  ||||| |||| |||||||||||| ||| |||||||||| |||||||
Db   307  GGACAAGTACAGGCTGCACCTGCCACCTGTGTGGCCGACCTCTTTGTCATCACGCTTCCCTT  366

Qy   404  CTGGGCAGTTGATGCCATGGCCGTCGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA  463
          |||||||||| |||||| ||  ||  ||||||| || |||||||||| |||||||||||
Db   367  CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGACTTTGGGAACTTCCTATGCAAGGCAGTCCA  426

Qy   464  TATCATCTACACTGTCAACCTCACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGA  523
          ||||||||||  ||| |||||||||||||| ||  | |||||||| ||||| ||||||
Db   427  TGTCATCTACACAGTCAACCTCTACAGCAGCAGTGTCCTCTCCTGGCCTTCATCATCAGTCTGGA  486

Qy   524  CCGGTACCTCCGCCATTGTCACGGCCAAGGCCAAGGAAACTGCTCGGCTGA  583
          ||| ||| || |||||||||| |||||||||||| | |||| | |||| ||
Db   487  CCGCTACCTCCGCCATCGTCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA  546

Qy   584  AAAGGCAGTCTATGTGGGCGTCTGGATCCCCAGCCCTCCTGCTATACCTGACTTCAT  643
          |||||||||  |||||||| |||||||| |||||| |||| | ||||| |||||||||
Db   547  AAAGGTGGTCTATGTGGGCGTCTGGGATCCCCTGCTATTCCCGTGATTCCCGACTTCAT  606

Qy   644  CTTTGCCGACGTCAGCCAGGCGGACTACAGTGACGGGGGACATCAGTTCAGGGGGAT  703
          ||||||||| |   |  |   ||||||| || ||||| ||||| || |   ||
Db   607  CTTTGCCAACG----------TCAGTGAGCAGATGACAGATATTATCTGTGACCG  651

Qy   704  CCTTTACCCCGATAGCCTGTTGGATGGTGGTCGTTTTCAATTCCAGCATATAATGTGGTGGTCT  763
          |||| ||||| |||||||||||||| | |||||||| || |||||| |||||| || |||||||
Db   652  CTTCTACCCCAATGACTTGTTGGGGTGGTTGTGTTCAGTTTCCAGTTTCAGCACCATCATGTGGCCT  711

Qy   764  CATCCTGCCCGGCATCGTCATCCTCTTCTGTTACTGCATCATCATCCTCTAAAGCTGTCACA  823
          ||||| |||| ||||||||| ||| ||  | ||  |||||||| |||| |||||| ||||||
Db   712  TATCCTGCCTGGTATTGTCATCCTGCTGTTGCTATTGCATTATCATCTCCAAGCTGTCACA  771

Qy   824  CTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACACAGTCAGTCATCCTCTAGCTTT  883
          |||||||| |||||||||||| ||||||| |||||||||||| |||||||| |||||||
Db   772  CTCCAAGGGGCCACCAGAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCATCCTCTGGCTTT  831

Qy   884  CTTTGCCTGCTGGCTCCCATTATGTGGGGATCCATGGGGATCAGTCGACTCCCTTTTGGG  943
          |||  |||| ||  ||||| ||| ||||| |||  || ||||||| ||| |||| |||||
Db   832  CTTCGCCTGCTGTTGGCTGCCTTACTACATGGGGATCCATGGGAACTCCACAGATTCCCTGGA  891

Qy   944  AGTCATCAAGCAAGATGATCTTCGAGAGCATTGTCACAAGTGGATCTTCCATCACACAGA  1003
          ||||||||| ||||| |  |||  |  | ||| | |||  |||| || ||||| |||||||
Db   892  AATCATCAAGCAGACAGAGGGTGTGAGTTTGAGAACCTGTGCCAAGCAAGTGGATTCCATCACCGA  951

Qy  1004  GGCCCTGCTTCTTCCACTGTTGCCTGAACCCCATCCTCTTCTCCTCGGGGCCAA  1063
          ||||||||| ||||||| |||| || |||||||||||||| ||| ||| |||||
Db   952  GGCCCTAGCTTCTTCCACTGTTGCTTGTCTATGCCTCATGCTTTGTTCCTTGGAGCCAA  1011

Qy  1064  GTTCAAAAGCTCTGCCCAGCCATGCACTCAAGCACCTCACCTTCTGTGAGCAGAGGGAAGAAA  1123
          ||  || |||||||||||| ||||||| |||| | |||||  |||||| |||||||||||||
Db  1012  ATTTAAAACCTCTGCCCAGCCACCACTCAACCCTTCACCTTGTGTGAGCAGAGGGTCCAGCTCAAGAT  1071
```

```
Qy  1124  CCTTTCCAAAGAAAGGGGGGGTGGACACTCTTCCGTCTCCACGAGTCAGAATCCTCCAG  1183
          ||| |||||||||||| |  ||||| | ||||| || |||||| || | |||| || |||
Db  1072  CCTCTCCAAAGAAAGGGGGAGGTGGGACATTCATCATCGTGTTTCCACTGTCTCTGAGTCTTCAAG  1131

Qy  1184  TTTTCACTCCAGCTAACCTTATGCAAAGACTTATATAATAATAATATGATAAA  1243
          ||||| | | ||| ||| ||||  |||| | | |||||| | |||||||
Db  1132  TTTTCACTCCAGCTAACACAGATGTAAAAGACTTTT------TTTATCGATAAATA  1183

Qy  1244  GAACTTTTATGTTACACATTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT  1303
          |||||||||| |||| |||||| |||||||| |  | || ||||||||||||||||
Db  1184  ACTTTTTTTAAGTTACACATTTTCAGATATAAAAGAACTGACCAATATTGTACAGTTTT  1243

Qy  1304  TTTTTTTTTTTAATTGACTGTGTGGGAGTTATGTCCTCTAGTTTTTGTGAGGTTTGACT  1363
          |||| |||  |||||||  ||||||| |  | |
Db  1244  TATTGCTTGTTGGATTTTTGT-------CTTGTGTTCTTTTAGTTTTTGTGAAGTTTAATT  1297

Qy  1364  TAATTATATAAATATTGTTTTTTGTTTCATGTGAATGAGCGTCTAGGCAGCAGGACCT  1423
          |    |||| |   |     |||||| |||| |||| |||| ||||||| ||||| |||
Db  1298  GACTTATTTATA----TAAATTTTTTGTTTCATATTGAATTGATGTGTCTTAGGCAGGACCT  1353

Qy  1424  GTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTAGGACTGTGTAGAGGAGA  1483
          |||||||||||||| |||||||| | || |||| ||||| |||||||||||
Db  1354  GTGGCCAAGTTCTTAGTTGCTGTTATGTCTCGTGGTAGGACTGTAGAA-------AAGGG  1405

Qy  1484  AACTGAACATTCCAGAATGTGTGTAAATTGAATAAAGCTAGCCCGTGATCCTCAGCTGTT  1543
          ||||||||||||||||| ||||||| |||| |||||||| || ||| || |||||||||||
Db  1406  AACTGAACATTCCAGAGCCGTGAGTAGTGAATCACGTAAGCTAGAAATGATCCCCAGCTGTT  1465

Qy  1544  GCTGCATA  1551
          | ||||||
Db  1466  TATGCCATA  1473
```

RESULT 9
US-10-225-567A-75
; Sequence 75, Application US/10225567A
; Publication No US2003011379BA1
; GENERAL INFORMATION:
;   APPLICANT: LifeSpan Biosciences
;   APPLICANT: Brown, Joseph P.
;   APPLICANT: Burmer, Glenna C.
;   APPLICANT: Roush, Christine L.
;   TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;   FILE REFERENCE: 1920-4-4
;   CURRENT APPLICATION NUMBER: US/10/225,567A
;   CURRENT FILING DATE: 2001-12-19
;   PRIOR APPLICATION NUMBER: 60/257,144
;   PRIOR FILING DATE: 2000-12-19
;   NUMBER OF SEQ ID NOS: 2292
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-75

Query Match   50.8%; Score 952.8; DB 15; Length 1679;
Best Local Similarity 80.2%; Pred. No. 5e-209;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

```
Qy    44  GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG  103
          ||||| |||||||||||||||  | ||||  ||||  ||||||||||| |||||| ||||
Db    12  GCGGCAGCAGCAGTAGCAAAGTGACCGCGAGGGCCTGAGTGCTCCAGTAGCACCGCATCTG  71

Qy   104  TAGAGGCGGAGTGTTGCCATGA----GGGGATCAGTATAACTACTCTG  163
Db    72  GAGAACCCAGCAGCGGTTACCATGA----GGGGATCAGTATATACACTTCAGATAACTACAC  126

Qy   164  TGAAGAAGTCGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAA  223
Db   127  CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGAACCCATGACCAATATTTGTGAAGAAA  186

Qy   224  CGTCCATTCAATAGGATCTTCCTGCCACCATCACTTCATCATCTTCTTGACTGGCAT  283
Db   187  TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTTCCTCCACCATCCTTCTTAACTGGCAT  246
```

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTGAGGCGGTTGGTGCTCCGGTAACCACCACGGCTG 103
              |||||| |||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   12 GCGGCGCAGGTAGCAAAGTGACGCGCGAGGGCCTGAGTGCTCCCAGTGACCCGCATCTG   71

Qy   104 TAGAGGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   72 GAGAACCAGGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 CGAGGAAATGGGGCTCAGGGGACTATGACTCCATGGAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy   224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTGACTGGCAT 283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 TGCTAATTTCAATAAAATCTTCGTGCCCACCATCTACTACTCCATCATCTTCTTAACTGCAT 246

Qy   284 AGTCGGCAATGGATTGGTGATCCTGGTCATCGGGGTTACCAGAAGAAGCTAAGGACATGAC 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 TGTGGGCAATGGATTGGTCATCCTGGTCATCGGGGTTACCAGAAGAAACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGTCACTTGTCATCACACTCCCCTT 403
              |||||||||||||||||||||||||||||||||||||||||||||
Db   307 GGACAAGTACAGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCATCATCACGTTCCCTT 366

Db   187 TGCTAATTTCAATAAAATCTTCCATCATCTTCTTAACTGGCAT 246
              |||||||||
Db   AGTCGGCAATGGATTGGTGATCCTGGTCATCGGGTTACCAGAAGAAGCTAAGGACATGAC 343
Qy   TGTGGGCAATGGATTGGTCATCCTGGTCATCGGGGTTACCAGAAGAAACTGAGAAGCATGAC 306
Db   GGACAAGTACCGGCTGCACCTGTCAGTGGCTGCAATTTTGTGTAGGCTGTCCA 463
Qy   GGACAAGTACAGGCTGCACCTGTCAGTGGCTGCCTATGCGAAGGCAGTCCA 426
Db   TATCATCTACACTGTCAACCTTACAGCAGCGCTTCTCATCCTGGCCTTCATCAGCCTGGA 523
Qy   TGTCATCTACACAGTCAACCTCACAGCAGTCAACCTGGCCTTCATCAGCCTGGA 486
Db   CCGGTACCTCGCCATTGTCCACGCCACCAAACAGTCAAGAGGCAAGAAACTGCTGGCTGA 583
Qy   CCGCTACCTGGCCATCGTCCACGCCACCAAACAGTCAAGAGGCAAGAAGCTGTTGGCTGA 546
Db   AAAGGCAGTCTATGTGGGCGTCTGGATCGGGGTTCCTCCTGACTATACCTGACTTCAT 643
Qy   AAAGGTGTCTATGTGGGCGTCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCAT 606
Db   CTTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGGGATGACAGTACATCTGTGACCG 703
Qy   CTTTGCCAACG-------TCAGTGAGGCAGTGACGATGACAGATATATCTGTGACCG 651
Db   CCTTTACCCCGATAGCCTGTGTGATGGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCT 763
Qy   CTTCTACCCCAATGACTTGTGGGTGGTGTTGTGTTCCAGTTCAGCACATCATGGTTGGCCTT 711
Db   CATCCTGCCCGGCATCGTCATCGTCTCCTCTGTTACTGCATCATCATCCTAAGGCTGTCACA 823
Qy   TATCCTGCTGGTATTGTCATCCTGCTATTGGCTCATTATCATCTCCAAGCTGTTCACA 771
Db   CTCCAAGGGCCACCAGAAGCGCAAGGCCCATATATGTGGGGATCAGCAGTCAGCATCAGCTTTT 883
Qy   CTCCAAGGGCCACCAGAGCGCAAGGCCCTCAAGACACCAGTCATCTCCTCATCCTGGCTTTT 831
Db   CTTTGCCTCTGCTGGCTGCTGCCCATATTATGTGGGCTGCTGCCTTACTCACATTGGGATCCTG 943
Qy   CTTCGCCTGTTGGCTGCTGCCTTACTACACATTGGGATCAGCATGCAGATCAGCATCCTGA 891
Db   AGTCATCAAGCAAGGATGTGACTTCGAGAGCAGCATTGTGCACAAGTGGATCTCCATCACAGA 1003
Qy   AATCATCAAGCAAGGTGTGAGTTGTGACAAGTGCATTGTGCACAAGTGGATTCCATCACCGA 951
Db   GGCCCTCGCCTTCTTCCACTCGTTGCCTGAACCCCATCCCATGCCTTCCTCCTGGGGGCCAA 1063
Qy   GGCCCTAGTTTTCTTCCACTCGTTGTCTGAACCCCATCCCTCTATGCTTTCCTTGGAGCCAA 1011
Db   GTTCAAAAGCTGTCTGCCAGCACTGGGGTGGACACTTCTTCCGTCTCCCAGCGAGGCAGAGAGGCCTCCAGGAT 1123
Qy   ATTTAAAACCTCTGCCCAGCACCACTTCCACCTCTGTGAGCAGAGGGTCCAGCCTCAAGAT 1071
Db   CCTTTCCAAAGGAAGAAAGCGGGGTGGACATTCCTTCCGTCCTTCCGGTCTTCCACCGGAGTCAGAATCCTCCAG 1183
Qy   CCTCTCCAAAGGAAGAAGCGAGGTGGACATTCCTCTGTTCCACTGAGTCGAGTCTGAGTGAAGAAA 1131
Db   TTTTTCACTTCAGCTAACCCTTATGCAAAGACTTATATAAAATATATATATGATAAA 1243
Qy   TTTTTCACTTCAGCTAACCCTTATGCAAAGACTTATATAATATAATATATATGATGATAAA 1183
Db   TTTTTCACTTCCAGCTAACACAGATGTAAAAGACTTTTT------TTTATACGATAAATA 1183
Qy   GAACTTTTTATGTTACACATTTTCCAGATATAGAGACTGACCAGTCTTGTACAGTTTT 1303
Db   ACTTTTTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTT 1243
Qy   TTTTTTTTTTAATTGACTGTTGGGGAGTTTATGTTCCTCTAGTTTTTGTGAGGTTTGACT 1363
Db   GGACAAGTACCGGCTGCACCTGTCACTTGTCATCACACTCCCCTT

Qy   1544 GCTGCATA 1551
            ||||||||
Db   1466 TATGCATA 1473

RESULT 11
US-10-341-434-177
; Sequence 177, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
;   APPLICANT: OriGene Technologies
;   TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
;   FILE REFERENCE: 9U 204 205 R1
;   CURRENT APPLICATION NUMBER: US/10/341,434
;   CURRENT FILING DATE: 2003-07-18
;   PRIOR APPLICATION NUMBER: US 60/348,164
;   PRIOR FILING DATE: 2002-01-15
;   PRIOR APPLICATION NUMBER: US 60/348,119
;   PRIOR FILING DATE: 2002-01-15
;   NUMBER OF SEQ ID NOS: 238
;   SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 177
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (89)..(1144)
;   OTHER INFORMATION:
US-10-341-434-177

Query Match   50.8%;   Score 952.8;   DB 16;   Length 1679;
Best Local Similarity   80.2%;   Pred. No. 5e-209;
Matches 1210;   Conservative   0;   Mismatches 252;   Indels 46;   Gaps 6;

Qy   44 GCAGTGCAGGTAGCAGTGACCCTCGAGCGTTGGTGCTCCGTAACCACCACGGCTG 103
            ||| || || || || || ||||| | ||| | | |||| ||| || ||| |||
Db   12 GCGGCAGCAGGTAGCAGTAGCAAAGTGACGCCGAGGCCTGAGTGCTCCAGTGCTCCACCGCATCTG 71

Qy   104 TAGAGCGAGTGTGTGCCATGGAACCGATCAGTGTGAGTATAACACTTCTGATAACTACTC 163
            |||||| ||||| ||||| || |||| || |||||||| |||||||||| |||| ||||||
Db   72 GAGAACCAGCGGTACCATGGA-----GGGGATCCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGGAGACATGAGCATCCAACAAGGAACCCTGCTTCGCGGGATGAAAA 223
            ||| ||||| |||||||| || ||||| ||||| ||| || |||| ||| ||||| ||||
Db   127 CGAGGAAATGGGCTCAGGGACTATGACTCCATGAGGAGAACCCTGTTTCCGTGAAGAAAA 186

Qy   224 CGTCCATTTCAATAGGATCTTCCTGCCACCATTACTTCATCATCTTCTTGACTCGGCAT 283
            ||| |||| |||| |||||| ||||| || ||| | ||| ||| ||||| |||| || |||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTCCCATCATCTTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGGATTGGTGATCCTGTCATGGTCATGGGTTACCAGAAGAGCTAAGGACATGAC 343
            |||| || || | ||||||| ||||| |||||| ||||| |||| || ||| || ||||| ||
Db   247 TGTGGGCAATGGATTGGTCATCCTGTCATGGGTTACCAGAAGAAGACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGGCTGGTACTTGGCTGGAAATTTTGTGTAAGGCTGTCCA 403
            |||| ||| ||||| || || || ||| ||||| |||| || |||| || ||| ||| ||
Db   307 GGACAAGTACAGGCTGCACCTGCACCTGGGGCCGACCCTCCTCTTTGTCATCACACTCCCCTT 366

Qy   404 CTGGGCAGTTGATGCCATGGCTGCACCTGGTACTTGGCTGGAAATTTTGTGTAAGGCTGTCCA 463
            |||| |||||| |||||| |||||||||||| || ||| ||| ||||||| |||||||| ||
Db   367 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCAACCTCTGGCCTTCATCAGCCTGA 523
            | |||||||| || || ||||| |||||||| |||| |||||| ||| |||||| ||||||
Db   427 TGTCATCTACACTGTCAACCTCTACAGCAGCGTCCTCAACCTCTGGCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTCGCCATTGTCCATTGTCCACAGTCAAGGCCAAGGAAACTGCTGGCTGA 583
            ||| ||||| || ||||||| |||||| || ||| |||||||||| || ||| ||
Db   524 GTGGCCAAGTTCTTAGTTGCTTAGTTCTTAGTTCTCGTGTAGGACTGTAGAA------AAGGG 1405

```
; Sequence 20, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
;  APPLICANT: Peterson, David P.
;  APPLICANT: Pearson, Cecelia I.
;  APPLICANT: Cocks, Benjamin G.
;  TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
;  FILE REFERENCE: PA-0029 US
;  CURRENT APPLICATION NUMBER: US/09/971,392
;  CURRENT FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 60/237,652
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 260
;  SOFTWARE: PERL Program
; SEQ ID NO 20
;  LENGTH: 1711
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Template ID: 284616.2
;  NAME/KEY: unsure
;  LOCATION: 1706
;  OTHER INFORMATION: a, t, c, g, or other
US-09-971-392-20

Query Match             50.8%;   Score 952.8;   DB 10;   Length 1711;
Best Local Similarity   80.2%;   Pred. No. 5e-209;
Matches 1210;  Conservative    0;  Mismatches 252;  Indels  46;  Gaps    6;
```

```
Qy       44 GCAGGTGCAGGTAGCAGTGGCCGTTGGTGCTCCGGTAACCACCACGGCTG 103
            ||||||||||||||||||||| |||||| |||| |||||||||||||||||
Db       47 GCGGCAGCAGGTAGCAAAGTGACGCGCGAGGGCCTGAGTGCTCCAGTCCTG 106

Qy      104 TAGAGCGGAGTGTTGCCATGGACCGATCAGTGTGAGTATATACACTTACTC 163
            ||||||||||||||||||||||||||    ||||||||||||||||||||||
Db      107 GAGAACCAGCGGGTTACCATGGA----GGGGATCAGTATATACACTTCAGATAACTACAC 161

Qy      164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACACAGGAACCCTGCTTCCGGGATGAAAA 223
            ||||||||||||||||||| |||||||| |||||||||||||||||||||||||||||||
Db      162 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 221

Qy      224 CGTCCATTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
            |||||||||||||||| ||||||||||||||||||||| ||||||| ||||||||||||
Db      222 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCCATCATCTTCTTAACTGGCAT 281

Qy      284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCATGAC 343
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      282 TGTGGGCAATGGATGGTGATGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGAC 341

Qy      344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACACTCCCCTT 403
            |||||||||||||||||||||||| ||||||||||||||||||
Db      342 GGACAAGTACAGGCTGCACCTGTCCAGTGGCCGACACTCCCTTTCTTCCCTT 401

Qy      404 CTGGGCAGTGATGCCATGGCTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
            ||||||||||||||||||| |||||||||||||||||||||| |||||||
Db      402 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCTCCATGCCATGGCAGTCCA 461

Qy      464 TATCATCTACACTGTCAACCTTCACAGCAGCCGTTCATCGGCCTTCATCCTCTGGA 523
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      462 TGTCATCTACACAGTCAACCTTCACAGCAGTCCTCATCCTGCCTCATCAGTCTCGA 521

Qy      524 CCGGTACCTCGCCATTGTCCACGCCGTGATCCGTGTT 583
            ||||||||||||||||||||||||||||||||||||||
Db      522 CCGGTACCTCGGCCATCGTCCACGCCGCCACCAACAGTCAGAGGCCAAGGAGCTGTTGGGCTGA 581

Qy      584 AAAGCAGTCTATGTGGGGTCTGGA 643
            ||||||||||||||||| |||||||
Db      582 AAAGGTGGTCTATGTTGGCGTTTGATCCTCGTGACTTCATTCCCGTGACTTCAT 641

Qy      644 CTTTTGCCGACGTCAGCCAGGGGGATGACAGATATCTGTGACCG 703
            |||||||||||
Db      642 CTTTGCCAACG------------------TCAGTGAGGCAGATGACAGATATCTGTGACCG 686
```

Query Match   50.8%; Score 952.8; DB 15; Length 1711;
Best Local Similarity   80.2%; Pred. No. 5e-209;
Matches 1210; Conservative   0; Mismatches 252; Indels 46; Gaps   6;

```
Qy     44  GCAGGTGCAGGTAGCAGTGACCCTCTGAGCGTTTGGTGTCTCGGTAACCACCACGGCTG  103
           |||||||||| ||| || || || ||||||||||||| |||||||||| || |||||||
Db     47  GCGGGCAGCAGGTAGCAAAGTGACCCGCGAGGGCCTGAGTGCTCCAGTAGCCACCCATCTG  106

Qy    104  TAGAGCGGAGTGTTGCCATGCAACCGATCAGTGTGAGTATATACACTTCGTGATAACTACTC  163
                                                |||||||||||||||||||||||
Db    107  GAGAACCAGCGGTTACCATGGA------GGGGATCAGTATATATACACTTCAGATAACTACAC  161

Qy    164  TGAAGAAGTGGGGTCTGAGACHATGACTCCAACAAGAACCCTGCTTCCGGGATGAAAA  223
                 || |||||||||||| ||||||||||||| |||||||||||||||||||||||||
Db    162  CGAGGAAATGGGCTCAGGGGACTCATGACCTCATGAAGAAGCCCTGTTTCCGTGAAGAAA  221

Qy    224  CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTGACTCGACTGGCAT  283
           ||||||||||||||||||||||||||||| |||||||| ||||||||||| ||||||||||
Db    222  TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTCCATATCTTCTTAACTGGCAT  281

Qy    284  AGTCGGCAATGGATTGGTGATCCTGGTCAGTGGCTCAGTGACCTCCTCTTTTGTCATCACACTCCCCTT  343
                                                                              
Db    282  TGTGGGCAATGGATTGGTCATCCTGGTCAGTGGCTCAGTGACCTCCTCTTTTGTCATCACACGCCTTCCCTT  341

Qy    344  GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACTCCAGTCTTGTGGGAAATTTTTGTGTAAGGCGTGTCCA  403
                                                                                 
Db    342  GGACAAGTACAGCGGCTGCACCTGTCAGTGGCGGCGGACCTCGTTTGTCATCGACTTGGGAACTTCCTATGCAAGGCAGTCCA  401

Qy    404  CTGGGCAGTGATGGCCATGCACGAGCTGCGTACTTGGGAAATTTTTGTGTAAGGCGTGTCCA  463
                                                                        
Db    402  CTGGGCAGTGATGGCCATGCAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA  461

Qy    464  TATCATCTACACGTCAACCTCACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGA  523
           |||||| ||||||||||||||| ||||||||||||||| |||||||||||| |||||||
Db    462  TGTCATCTCCACAGTCAACCTCACAGCAGCGGTGTCCTCATCCTGGCCTTCATCAGTCTGGA  521

Qy    524  CCGGTACCTGCCATTGTCAACGCCACCAACAGTCAAGGCCAAGGAAACTGCTCAGGCTGA  583
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522  CCGCTACCTGCCATCGTCAACGCCACCAACAGTCAAGGCCAAGGAGAAGCTGTTGGCTGA  581

Qy    584  AAAGGGCAGTCATGTGGGCGGTCTGAGGATCCCAGCCCCTCCTCCTGTGATCCTGACTACCTCAT  643
                                                                            
Db    582  AAAGGTGGTCTATGTGGGCGTCTGGATCCCGTGGGATCCCTGCCCCTCCTGCTCAGCAGTACCTCAT  641

Qy    644  CTTTGCCGAGTCAGCCAGGGGACTCAGTGGGGGACATCAGTGGGTACATCTGTGACCG  703
                                                              |||||||||
Db    642  TCAGTGAGGCAGATGACAGATATCGTGACCG  686

Qy    704  CCTTTACCCCGATAGCCTGTGGATGGTCGGATGGTGTTCAATTCCAGCATATAATGTGGGTCT  763
           |||||||||||                                                     
Db    687  CTTCTACCCCCAATGACTTGTGGGTGTTGGTTTCAGCACATCATGGTTGGCCT  746

Qy    764  CATCCTGCCCGGCATCGTCATCCTCGTGTTACTGCATCATCATGTCATTGCATTATCATTCCAAGCTGTCACA  823
                                                                                   
Db    747  TATCCTGCCTGGTAATTGTCATTGCATTGCTGCTATTGTCATTGCATTATCATTCCAAGCTGTCACA  806

Qy    824  CTCCAAGGGGCCACCAGAAGCCAAGCCCTCAAGACGACAGTCATCCTCATCCTAGCTTT  883
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    807  CTCCAAGGGGCCACCAGAAGCCAAGCCCTCAAGACCCTCAAGACCCACAGTCATCCTCATCCTGGCTTT  866

Qy    884  CTTTGCCTGTCTGGCTGCCATATTATGTGGGGATCAGCATCGACTCCTTCATCCTTTTGGG  943
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    867  CTTTGCCTGTCTGGCTGCCATTGCACTTGGGATCAGCATCGACTCCTTCATCCTTCCCTGA  926

Qy    944  AGTCATCAAGCAAGGATGTGACTTCGAGAGAGCATTTGCACAAGTGGATCTGCATCACAGA  1003
           | ||||||||||||||||||||||||||||| ||||||||||| ||| ||||||||||||
Db    927  AATCATCAAGCAAGGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGA  986

Qy   1004  GGCCCTGCCCTTCTTCCACTGTGCTGCACGACCCCATCCTCTATGCCTTCCTCGGGGCCAA  1063
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    987  GGCCCAGCTTTCTTCCACTGTGCTGCACCCCATCCTCTATGCCTTCCTCTGGAGCCAA  1046

Qy   1064  GTTCAAAAGTCTGCCCCAGCATGCACTCAACTCAACTGAGCAGAGGCTCCAGCCCTCCAAGAT  1123
           ||||||| | |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1047  ATTTAAAACCTCTGCCCAGCACCTCAACTCAACTGAGCAGAGGTCCAGCCCTCCAAGAT  1106

Qy   1124  CCTTTCCAAAGAAAGCAAGGGGGGTGGACACTCTTCCGGTCCCATCGAGAGTCAGAATCCTCCAG  1183
                                                                              
Db   1107  CCCTCCAAAGGAAAGCGAGGTGGACATTCATCCTCTGAGTCTGAGTCTTCAAG  1166

Qy   1184  TTTTCACTCCAGCTACCCTTATGCAAAGACTTATATAAATAATATAATATATATGATAAA  1243
                                                                           
Db   1167  TTTTCACTCCAGCTAACACAGATGTAAAAGACTTTTT------TTTATACGATAAATA  1218

Qy   1244  GAACTTTTTTTATGTTACACATTTTCCAGATATAAGAGAGACTGACCCAGTCTGTACAGTTTT  1303
                                                                          
Db   1219  ACTTTTTTTTTAAGTTACACATTTTCAGATATAAAAGACTGACCAATATTGTACAGTTTT  1278

Qy   1304  TTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCCTCCAGTTTTGTGAGGTTTGACT  1363
                                       |||||||||||||||||||||||||||||
Db   1279  TATTGCTTGGTTGTTGTTTT-----CTTGTGTTTTCTTTAGTTTTTGTGAGGTTTAATT  1332

Qy   1364  TAATTTATATAAATATATTGTTTTTTGTTTGTTCATGTGAATGAGCGTCTAGGCAGGACCT  1423
                                                                         
Db   1333  TAAATTTTTTTTGTTTTGTTCATATTGCTGTATGTGTGTGTCTGGTAGGCAGGACCT  1388

Qy   1424  GTGGCCAAGTTCTTAGTAGCTGTGTTATCTGTGTGTAGGAACTGTTAGAGGAAGA  1483
           |||||||||||||||||||||||||| |||||||||||||||||||||      
Db   1389  GTGGCCAAGTTCTTAGTTGCTGTATGTCTCGTGGTAGGAGGACTGTAGGA-------AAGGG  1440

Qy   1484  AACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTT  1543
           |||||||||||||||||||||||||||| |||||||| ||||||||||||||||| ||||
Db   1441  AACTGAACATTCCAGAGCGTGTAGTCAATCACGTAAAGCTAGAAATGATCCCCCAGCTGTT  1500

Qy   1544  GCTGCATA  1551
           ||||||||
Db   1501  TATGCATA  1508
```

RESULT 13
US-10-101-510-459
; Sequence 459, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20

```
Query Match     50.7%;   Score 951.8;  DB 17;   Length 1664;
Best Local Similarity   80.2%;   Pred. No. 8.4e-209;
Matches 1209;   Conservative   0;   Mismatches 252;   Indels  46;   Gaps  6;

Qy   45 CAGGTGCAGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGT 104
         | |||||||||||| || ||| ||  ||||  || || ||||| ||||| ||| ||
Db    1 CGGGCAGCAGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTGCCACCGCATCTGG  60

Qy  105 AGAGCGAGTGTTGCCATGGAACCCATCAGTGTCAGTATAACACTTCGATAACTACTCT 164
         ||| | ||| || |||||||    || |||||||  || ||| ||| |||| ||| ||
Db   61 AGAACCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACC 115

Qy  165 GAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCGCTTCCGGGATGAAAAC 224
         || ||| |||| || ||||||||||||||||  |||||||||| |||||| |||| ||
Db  116 GAGGAAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCGTGTTTCCGTGAAGAAAAT 175

Qy  225 GTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTCCATCACTGGCATA 284
         | | ||||||||| | |||||||||||||||||||||  ||||||||| |  ||||||
Db  176 GCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTCCATCATCTCCTTAACTGGCATT 235

Qy  285 GTCGGCAATGGATTGGTGATCCTGGTCAGGGTTACCAGAAGAAGCTAAGGAGCATGACG 344
         || |||||||||||||| ||||||||||| |||||||||||||||| |||||||||||
Db  236 GTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGCTGAGGAGCATGACG 295

Qy  345 GACAAGTACCGGCTGCACCTGTCAGTGCTGACTGCTCAGTGGCTGACCTCCTCTTGTCAT 404
         ||||||||| || ||||||||  || || |||||||||| |  |  | ||||||||||
Db  296 GACAAGTACAGGGTGCACCTGCCACCTGTCAGTGCCAGGCCTGAGTGCTTCCCTTC 355

Qy  405 TGGGCAGTGATGCCATGGCATGTGGTACTTGGGAAATTTGTGTAAGGCTGTCCAT 464
         ||| |||||| |||| ||  ||| |||||||||||  ||  |||||| |||||||
Db  356 TGGGCAGTGATGCCGTGCAACTGGTACTTGGGACTTCCTATGCAAGGCAGTCAT 415

Qy  465 ATCATCTACACTGTCAACCTCACACAGCAGGTGTTCTCATCTGCCTTCATCAGCCTGGAC 524
         || ||||||||| |||||||||||||||||| |  | || |  || | | |||||||
Db  416 GTCATCTACACAGTCAACCTCACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGAC 475

Qy  525 CGGTACCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAA 584
         || ||||| || |||| |||||| ||||||||| |||||||||||||| ||||||||
Db  476 CGCTACCTGGCCATCGTCCAGCGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAA 535

Qy  585 AAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCCTCCTGACTATCACCTGACTTCATC 644
         |||||||||||| |||| |  || |  |||  |||||||||| ||||  | ||||||||
Db  536 AAGGTGGTCTATGTTGGCGTTCTTGGATCCTGGATCCCAGTCCCTGCTGACTATTCCGACTTCATC 595

Qy  645 TTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGC 704
         |||||||||||                  |||||||| ||| | | |||| ||||||||
Db  596 TTTGCCACG----------------TCAGTGAGGCAGCAGATGACGCAGATATCTGTGACCGC 640
```

```
Qy  884 CTTTGCTGCTGGCTGCCATATTATGTGGGATCAGCATCGACTCCTTCATCCTTTGGG 943
         ||| || || ||||||||     | |||||||||||||||| ||||||||||| ||
Db  867 CTTCGCCTGTTGGCTGCCTTACACATTGGGATCAGCATCGACTTCCTTCATCCTCCTGA 926

Qy  944 AGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACAGA 1003
         | ||||||||||||  ||||| || ||||  ||||||||||||||||||| |||||| ||
Db  927 AATCATCAAGCAAGGGGTGTGAGTTGGAGAACACTGTGCACAAGTGGATTTCCATCACCGA 986

Qy 1004 GGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCATCCTCTATGCCTTCCTCCTGGGGCCAA 1063
         |||||| || |||||||||||||| |||||||| |||||||||| |||||  | ||||||
Db  987 GGCCCTAGCTTCTTCTTCCACTGTTGTCTGAACCCCATCCTCTATGCTTTCCTTGGAGCCAA 1046

Qy 1064 GTTCAAAAGTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGAT 1123
         ||||||||  |||||||||| ||||||| || |||| ||||||| ||||||||||||
Db 1047 ATTTAAAACCTCTGCCCAGCACGCACTCACCTCTGTGACCAGATCCCAGCCTCAAGAT 1106

Qy 1124 CCTTTCCAAAGGAAAGCGGGGTGGACACCTCCTCCGTCTCCACCGAGTCAGAATCCTCCAG 1183
         ||| |||||||||||| | |||||||| | | ||||| ||||| |||||| |  ||||| 
Db 1107 CCTCTCCAAGGAAAGCAGCGGAGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAG 1166

Qy 1184 TTTTCACTCCAGCTAACCCTTATGCAAAGACTTATATAATATATAATATATATGATAAA 1243
         |||||||||||||||||   | |||||||||||| ||   |  ||   | ||  |||| 
Db 1167 TTTTCACTCCAGCTAACACAGATGTAAAAAGACTTTTT------TTTATACGATAAATA 1218

Qy 1244 GAAACTTTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAGTCTTGTACAGTTTT 1303
         |  ||||||||  |||||||||||| |||||||| ||||||| ||||  |||||||||||
Db 1219 ACTTTTTTTTTAAGTTACACATTTTTCAGATATAAAAGACGACCAAATATTGTACAGTTTT 1278

Qy 1304 TTTTTTTTTTTTAATTGACTGTTGGGAGTTTAATGTTCCTCTAGTTTTTGTGAGGTTTGACT 1363
         | ||| ||||| ||| ||||| |||  ||||              | ||||||| || || 
Db 1279 TATTGCTTGTTGGATTTTGT------CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAATT 1332

Qy 1364 TAATTATATAAAATATTGTTTTTTTTTGTTTTCATGTGAATGAGCGTCTAGGCAGGACCT 1423
         |||||| | |||||| |||||||||||||| |||| | |||| || |||||||||||||
Db 1333 GACTTATTTATA-----TAAATTTTTTTGTTCATATTGATGTGTGTCTAGGCAGGACCT 1388

Qy 1424 GTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAGGACTGTAGAACTGTAGAGGAAGA 1483
         |||||||||||||||| |||||  || | ||||| ||||||||||||        ||| 
Db 1389 GTGGCCAAGTTCTTAGTTGCTGTGTATGTTCGTGTGGTAGGACTGTAGAA-------AAGGG 1440

Qy 1484 AACTGAACATTCCAGAATGTGTGGTAAATTGAATAAAGTAGCCGTGATCCTCAGCTGTT 1543
         |||||||||||||||| ||||| ||| |  |||  |||||| | |||||| ||||||||
Db 1441 AACTGAACATTCCAGAGCGTGTAGTGAATCACGTAAAGTAGAAATGATCCCCAGCTGTT 1500

Qy 1544 GCTGCATA 1551
         | |||||| 
Db 1501 TATGCATA 1508
```

```
;    CURRENT APPLICATION NUMBER: US/09/880,107
;    CURRENT FILING DATE: 2001-06-14
;    PRIOR APPLICATION NUMBER: US 60/211,379
;    PRIOR FILING DATE: 2000-06-14
;    PRIOR APPLICATION NUMBER: US 60/237,054
;    PRIOR FILING DATE: 2000-10-02
;    SOFTWARE: PatentIn Ver. 2.1
;    NUMBER OF SEQ ID NOS: 3950
; SEQ ID NO 2143
;    LENGTH: 1670
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: Genbank Accession No. US20020142981A1 L06797
US-09-880-107-2143

Query Match       50.3%;    Score 944.8;    DB 9;    Length 1670;
Best Local Similarity   75.9%;   Pred. No. 3.5e-207;
Matches 1356;   Conservative    0;   Mismatches 297;   Indels 133;   Gaps 10;
```

```
Qy     66 CTCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCGTGTAGAGGAGTGTTGCCATGGAA 125
                    |||||||  |||||||||   |   | ||| ||||||| ||||||||||||
Db     18 CGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAACCAGCGGTTACCATGGA 77

Qy    126 CCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGAGAC 185
                 |   ||   | | |||||||||   |||||  ||| |||||||||||
Db     78 -----GGGGATCAGTATATACACTTCAGATAACTACTCACCGAGGAAATGGGCTCAGGGGAC 132

Qy    186 TATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTC 245
                 |||||||||||||||||||| ||||||  ||||||  ||||| |||||||  |||| 
Db    133 TATGACTCCATGGAAGCAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTC 192

Qy    246 CTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATGGTCGGCAATGGATTGGTGATC 305
                 ||||||||||||||||||||||||  ||||||  ||||||||  ||||||||||| 
Db    193 CTGCCCACCATCTCCATCATCTTCTTACTGGCATTGTGGGCAATGGATTGGTCATC 252

Qy    306 CTGGTCATGGGTACCAGAGAAGCTAAGAGCCATGACGGACAAGTACCGGCTGCACCTG 365
                 |||||||  |||||||||||||||||||  ||||||  |||||||  |||||||
Db    253 CTGGTCATGGGTACCAGAGAAGAAACTGAGAGACTGAGAGCATGACGGACAAGTGCACCTG 312

Qy    366 TCAGTGGCTGACCTGCCTCCTTGTCATCACACTCCCCTCTGGGCAGTTGATGCCATGGCT 425
                 ||||||||||||||||||||||||||||||||||||  |||||||||||| ||| 
Db    313 TCAGTGGCCGACCTCCTTGTCATCAGCTTCCCCTTCTGGGAGTTGATGCCGTGGCA 372

Qy    426 GACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTC 485
                 || |||||||||||  ||||| |||||  || |||| ||| ||||||||||||| 
Db    373 AACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGCTCATCTACACAGTCAACCTC 432

Qy    486 TACAGAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCCAC 545
                 |||||||||||| ||||||||  |||||  ||||||||||||||||  ||||||||||
Db    433 TACAGAGTGTCCTCATCCTGGCCTTCATCAGTGTGGCTGGCATCGTCCAC 492

Qy    546 GCCACCAACAGTCAAAGGCCAAGGAAACTGTGGCTGAAAAGGCAGTCTATGTGGCGTC 605
                 ||||||||||||  |||||||||||||||||||||||||||||||  || |||||||||
Db    493 GCCACCAACAGTCAGAGGGCCAAGGAAGCGTGGCTGGCTGGAAAAGGTGGTCATGTTGGCGTC 552

Qy    606 TGGATCCCAGGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGG 665
                 |||||||||||||  |||||||||||||||| |||||||||||||||||
Db    553 TGGATCCCTGCCCCTCCTGCTGACTATTCCCGACTATTCATCTTTGCCCAACG---------- 601

Qy    666 GACATCAGTCAGGGGATGACGAAGCAGTTACATCTGTGACGCTGTTACCCCGATAGCCTGTCG 725
                               | |||||||||| | |||| |||| ||| |||||||||| | |||
Db    602 ----TCAGTGAGGCAGCAGATATAATCGTGACGATAATGACCGGCTTCTCACCCCAATGACTTGTCG 657

Qy    726 ATGGTGGTGTTTCAATTCCAGCATATAAATGGTGGGTCTCATCGTCGCCGGCCATCGTCATC 785
                 || |||| | ||||||||||||   || |  |||  ||| || |||||||||| 
Db    658 GTGGTTGTGTTGCCAGTTCAGCACACATCATGGTTGCCTTATCCATCCCGGCCTTATCCTTGTGTCATC 717

Qy    786 CTCTCCTGTTACTGCATCATCATCCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGGGC 845
                  ||| | ||||||||||||||||||| ||||||||||||||||||  |||||||||||||
Db    718 CTGTCCTGCTATTGCATTATCATCCTCCAAGCTGTCACACTCCAAGGGCCACCAAGAAGGC 777

Qy    846 AAGGCCCTCAAGACGACAGTCATCATCCTCATCCTAGCTTTCTTTGCCTGCTGGCTGCCCATAT 905
```

```
Qy    705 CTTTACCCCGATAGCCTGTGGATGGTGTTTCAATTCCAGCATATAATGTGGGGTCTC 764
                 ||||  |||||  |||||  ||||| |||||||||||||||||  |||| ||
Db    641 TTCTCACCCCAATGACTGTGGGTGGGTGGTTGTGTTGTTCCAGTTTCAGCACACATGTTGGCCTT 700

Qy    765 ATCCTGCCCGGCATCGTCATCCTCCTCGTTACTGCATCATCATCTCTAAGCTGTCACAC 824
                 |||||||  |||||||  ||||||||||  ||||  ||||| ||||||||||||||
Db    701 ATCCTGCTGGTATTGTCATCCTGTCCTGCTATTGCATTATCATCATCCTCCAAGCTGTCACAC 760

Qy    825 TCCAAGGGCCACCAGAAGGCCAAGGCCCTCAAGACGACAGTCATCATCCTCATCCTCAGCTTTC 884
                 |||||||||||||||||  |||||||||||||||||||||||||  |||||||||||||
Db    761 TCCAAGGGCCACCAACAGGCCAAGGCCCTCAAGACCACCAGTCATCATCCTCATCCTCAGCTTTC 820

Qy    885 TTTGCCTGCTGGCTGCCCATATATGTGGGGATCAGCAATGCATCAGCATCGTCATCCTTCATCCTTTTGGGA 944
                 |  ||| || ||||||||||| |||||||||| | |||  | |||||||||| ||  |||  ||
Db    821 TTCGCCTGTTGGCTGCCCTTACTACATTGGGATCAGCTAGCAGTTTCATCCTTCATCACCGAG 880

Qy    945 GTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCCTCCATCACAGAG 1004
                 ||||||  |||||||||| |||  ||||||||  |  || ||| ||| ||||| |
Db    881 ATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGAG 940

Qy   1005 GCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTTCCTCGGGGCCAAG 1064
                 ||||  |||||||||||||||||||||  |||||||||||  |||  |||||  |||||
Db    941 GCCCTAGTTTCTTCCCACTGTTGCTGAACCCCATCCTCTATGCTTTCCTTGGAGCCAAA 1000

Qy   1065 TTCAAAAGCTCTGCCCAGCATGCACTCACTCAACTCCATGAGCAGAGGCTTCCAGCCTCAAGATC 1124
                 |||||  |||||| |||| || |||||||  |||||| |||||||||  |||| ||| ||||
Db   1001 TTTAAAACCTCTGCCCAGCACGCACTCACCTGTGAGCAGAGGGGTCCAGCCTCAAGATC 1060

Qy   1125 CTTTCCAAAGGAAAGCGGGGTGACACTCTTCGTCTCCACGGAGTCAGAATCCTCCAGT 1184
                 |||||||||||||||  ||||||||||||||||||  || |||  ||| ||||||  
Db   1061 CTCTCCAAAGGAAAGCGGGTGACATTCATCTGTTTCCACTGTGTCTTCAAGT 1120

Qy   1185 TTTCACTTCCAGCTAACCCTTATGCAAAGACTTATATATTATATATATATGATAAAG 1244
                 |||||||||||||||||||||  ||||| |||||||||| |||||||||
Db   1121 TTTCACTTCCAGCTAACACAGATGTAAAAGACTTTTT---------TTTATACGATAAATAA 1172

Qy   1245 AACTTTTTATGTTACACATTTTCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTT 1304
                 |||| |||||||  ||||||||||||||| ||| |||  |||||||||||  ||||||
Db   1173 CTTTTTTTTAAGTTACACATTTTTCAGATATATAAAAGACTGACCAATATTGTACAGTTTTT 1232

Qy   1305 TTTTTTTTTTAATTGACTGTTGGGAGTTTATGTCCTCTAGTTTTTGTGAGGTTTGACTT 1364
                                   |||||||  ||| || ||  |||||||||||  ||| |||
Db   1233 ATTGCTTGTTGGATTTTTGT------CTTGTGTTTCTTTTAGTTTTTGTGAAGTTTTAATTG 1286

Qy   1365 AATTTATAAATATATGTTTTTTTGTTGTTTCATGTGAATGAGGCTCTAGGCAGGACCTG 1424
                  |  ||||  |||| | ||||||||  |||| |||||||||||||  |||  |||||||
Db   1287 ACTTATTTATA----TAAATTTTTTTTGTTTCATATTGATGTGTGTCTAGGCAGGACCTG 1342

Qy   1425 TGGCCAAGTTCTTAGTAGCTGTTTATTCTGTGTAGGACTGTAGAACTGTAGAGGAAGAA 1484
                 ||||||||||||||||||||| |||| |||||||||||  |||||||||||
Db   1343 TGGCCAAGTTCTTAGTTGCTGTAATGTCTGTAGGACTGTAGAA-------AAGGGA 1394

Qy   1485 ACTGAACATTCAGAATGTGTGTAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTTG 1544
                 ||||||||||||||| || || ||| ||| |||||||  ||||| ||||||||||||
Db   1395 ACTGAACATTCCAGAGACGTGTTAATCACGTAAAGCTAGAAATGATCCCCAGCTGTTT 1454

Qy   1545 CTGCATA 1551
                 |||||||
Db   1455 ATGCATA 1461
```

```
RESULT 15
US-09-880-107-2143
; Sequence 2143, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;    APPLICANT: Horne, Darci T.
;    APPLICANT: Vockley, Joseph G.
;    APPLICANT: Scherf, Uwe
;    APPLICANT: Gene Logic, Inc.
;    TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;    FILE REFERENCE: 44921-5028-WO
```

```
Db  778  ||||||||||| ||| |||||||||||||||||| |||||| | ||||||||||||||| ||||||||  837
         AAGGCCCTCAAGACCACAGTCATCCTCATCCTGGCTTCTTCGCCTGTTTGCTTGGCTGCCTTAC

Qy  906  TATGTGGGGATCAGCATCAGCATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGAC  965
         || |||| |||||||||||| |||||||||||||||| ||||| ||||||||||||||| |||||
Db  838  TACATTGGGATCAGCATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAG  897

Qy  966  TTCGAGAGCATTGTGCACAAGTGGATCCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGT  1025
         |||||||| ||||||||||||||||||  ||||||| ||||||||| || ||||||||||||
Db  898  TTTGAGAACACTGTGCCACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTCTTCCACTGT  957

Qy  1026 TGCCTGAACCCCATCCTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCAT  1085
         || ||||||||||||||||||| |||||| || |||||| ||||| ||||||||||||||
Db  958  TGTCTGAACCCCATCCTCTATGCTTTCCTTGGAGCCAAATTTAAAACCTCTGCCCAGCAC  1017

Qy  1086 GCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGGT  1145
         |||||| ||| | |||||||||| |||||||||||||||| |||||||||||||  |||
Db  1018 GCACTCACCTCTGTGAGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCCAGGT  1077

Qy  1146 GGACACTCCTTCCGCTCTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTA  1205
         ||||| | | | || || |||| ||| |||||||||||||||||||||||||||       
Db  1078 GGACAATCATCTGTTTCCACTGCTCGAGTCTTCAAGTTTCACTCCAGCTAA------  1131

Qy  1206 TGCAAAGACTATATATATATATATATATATATGATAAAGAACTTTTTATGTTACACATT  1265
         ||| |||| ||||||||||  |||  |||||| |||||| |||||||| |||||||||||
Db  1132 --CACAGATGTAAGAGACTTTTTTATACGATAAATAAACTTTTTTTAAGTTACACATT  1189

Qy  1266 TTCCAGATATAAGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTTAATTGACTGTT  1325
         |||||||||| |||||||||||  | ||||||||||||| || |   ||| ||| |||
Db  1190 TTTCAGATATAAAAGACTGACCAATATTGTACAGTTTTTTATTGCTT----GTTGGATTTT  1245

Qy  1326 GGGAGTTTATGTTCCCTCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAAATATTGTTTT  1385
         || | ||||||||| || |||||||| |||| ||||| || ||| ||||||||||| ||
Db  1246 TGCTCTTTGTGTTTCTTTAGTTTTTCGTGAAGGTTAATTGACTTATTTATATA---AATT  1302

Qy  1386 TTGTTTGTTTCATGTGAATGACGGTCCAGGCGACCTGTGGCCAAGTTCTTAGTAGCTG  1445
         || ||||||||||| | ||| | || ||||| |||||||||||||||||||||| ||||
Db  1303 TTTTTTGTTTCATATTGATGTGTGTCTCAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTG  1362

Qy  1446 TTTATCTGTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACATTCCAGAATGTGT  1505
         | ||| | |||| |||||| |||||| |      |||| ||||||||||||  ||
Db  1363 TATGTCTCGTGGTAGGACGTGTAGAA------AAGGGAACTGAACATTCCAGAGCGGTGT  1414

Qy  1506 GGTAAATTGAATAAAGCTAGCCGTGATCCTGCAGCTGTTGCTCGCATAATCTCTTCATTCCG  1565
         || |||||||| ||||||||   |||| | |||||||| | |||||| |||||||| ||
Db  1415 AGTGAATCACGTAAAGCTAGAGAAATGATCCCCAGCTGTTTATGCATAGATAATCTCTCCA-  1473

Qy  1566 AGGAGCACCCCACCCCCACCCCCCACCCCCATTCCTTAAATTGTTTGGTTATGCTGT  1625
         ------------------------------------------------------------
Db  1474 --------------------------------------------------------------  1473

Qy  1626 GTGATGGTTTGTTTGGTTTTTTTGTGTTGTGTTTGTTTTTTTCTGTAAAAGATG  1685
         || ||||       |||||||||||||| ||  |||| |||||||||||||||
Db  1474 ------TTCCCGTGGAACGTTTTTCCTGTTCTTAAGACGTGATTTTGCTGTAGAAGATG  1526

Qy  1686 GCACTTAAAACCAAAGCCTGAAATGGTGGTAGAAATGCTGGGGTTTTTTTGTTTTGTTTG  1745
         ||||||| || |||||||  |||| || |||||||||||||            
Db  1527 GCACTTATAACCAAAGCCCAAAGTGGT-ATAGAAATGCTGG------------  1566

Qy  1746 TTTTTCAGTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTGTATTA  1805
         ||||||||||||| |||||| ||| ||||||| | ||| |||||||||||||||||||
Db  1567 -TTTTTCAGTTTCAGGAGTGGGTTGATTTCAGCACTAC-AGTGTACAGTCTTGTATTA  1624

Qy  1806 CATTGTTAATAAAGTCAATGATAAACTTAAAAAAAAAAAAAA  1851
         ||||||||||||| | ||| ||||||||||||||||||||||
Db  1625 AGTTGTTAATAAAAGTACATGTTAAACTTAAAAAAAAAAAAA  1670
```

Search completed: May 17, 2004, 22:14:00
Job time : 934.766 secs

```
OM nucleic - nucleic search, using sw model

                    GenCore version 5.1.6
              Copyright (c) 1993 - 2004  Compugen Ltd.

Run on:     May 17, 2004, 05:50:31 ; Search time 5417.72 Seconds
                                      (without alignments)
                                      10345.919 Million cell updates/sec

Title:           US-09-367-052-1
Perfect score:   1877
Sequence:        1 ccatcctaatacgactcact.......aaaaaaaaaaaaaaaaaaaa 1877
Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched:        27513289 segs, 14931090276 residues
Total number of hits satisfying chosen parameters:   55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  EST:*
        1:  em_estba:*
        2:  em_esthum:*
        3:  em_estin:*
        4:  em_estmu:*
        5:  em_estov:*
        6:  em_estpl:*
        7:  em_estro:*
        8:  em_htc:*
        9:  gb_est1:*
       10:  gb_est2:*
       11:  gb_htc:*
       12:  gb_est3:*
       13:  gb_est4:*
       14:  gb_est5:*
       15:  em_estfun:*
       16:  em_estom:*
       17:  em_estfun_hum:*
       18:  em_gss_inv:*
       19:  em_gss_pln:*
       20:  em_gss_vrt:*
       21:  em_gss_fun:*
       22:  em_gss_mam:*
       23:  em_gss_mus:*
       24:  em_gss_pro:*
       25:  em_gss_rod:*
       26:  em_gss_phg:*
       27:  em_gss_vrl:*
       28:  gb_gss1:*
       29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	747.8	39.8	906	13	BU523025	BU523025 AGENCOURT
2	720.4	38.4	1091	13	BX397685	BX397685 BX397685
3	707	37.7	1201	13	BX421287	BX421287 BX421287
4	705.6	37.6	1201	13	BX396365	BX396365 BX396365
5	704.6	37.5	1201	13	BX421511	BX421511 BX421511
6	701	37.3	1201	13	BX358346	BX358346 BX358346
7	685	36.5	1201	13	BX353253	BX353253 BX353253
8	684.8	36.5	1051	12	BM545259	BM545259 AGENCOURT
9	660.8	35.2	1201	13	BX462627	BX462627 BX462627
10	647.4	34.5	1201	13	BX418530	BX418530 BX418530
11	631.4	33.6	958	12	BG174412	BG174412 60234232
12	626.6	33.4	1201	13	BX462182	BX462182 BX462182
13	621.6	33.1	793	9	AA182270	AA182270 mt83f01.r
14	621.2	33.1	639	9	AA816049	AA816049 vr14h02.r
15	621.2	33.1	952	13	BX362477	BX362477 BX362477
16	610.2	32.5	753	12	BG915636	BG915636 60814106
17	602.2	32.1	872	14	CD251331	CD251331 AGENCOURT
18	599.2	31.9	945	13	BX443258	BX443258 BX443258
19	594	31.6	910	12	BI762229	BI762229 603049139
20	593.4	31.6	911	13	BX443334	BX443334 BX443334
21	590	31.4	955	12	BG173867	BG173867 60233928
22	578.4	30.8	891	12	BI824663	BI824663 603033815
23	576.8	30.7	950	12	BQ718617	BQ718617 AGENCOURT
24	576.2	30.7	628	10	AW227957	AW227957 up19e03.y
25	576	30.7	904	10	BF100790	BF100790 601753725
26	574.2	30.6	886	13	BX443228	BX443228 BX443228
27	572	30.5	600	13	BU919690	BU919690 6033-58 M
28	566.2	30.2	928	14	CD389217	CD389217 AGENCOURT
29	560.6	29.9	581	14	CF540844	CF540844 UI-M-GW0-
30	558	29.7	1034	12	BM920800	BM920800 AGENCOURT
31	552.8	29.5	765	12	BM387369	BM387369 UI-R-CN1-
32	551.6	29.4	582	12	BG145042	BG145042 ut75a01.y
33	550.2	29.3	876	14	CA454255	CA454255 AGENCOURT
34 c	549.6	29.3	749	14	CD466324	CD466324 LeukoN2_3
35	549	29.2	1099	13	BX397684	BX397684 BX397684
36 c	547.4	29.2	827	12	BI761664	BI761664 603046395
37 c	539.4	28.7	1201	13	BX358340	BX358340 BX358340
38 c	532.4	28.4	1201	13	BX421141	BX421141 BX421141
39	524.8	28.0	528	12	BG277031	BG277031 uv10e03.y
40	524.8	28.0	750	12	BG613352	BG613352 602641234
41	520.4	27.7	875	12	BI756157	BI756157 603030061
42	519.2	27.7	976	13	BU146456	BU146456 AGENCOURT
43	511.4	27.2	585	10	BE627479	BE627479 uu52b07.y
44	509.6	27.1	913	12	BM051973	BM051973 603638877
45	509	27.1	886	10	BF101953	BF101953 601752919

ALIGNMENTS

```
RESULT 1
BU523025
LOCUS       BU523025      906 bp    mRNA    linear    EST 13-SEP-2002
DEFINITION  AGENCOURT_10154335 NCI_CGAP_Co24 Mus musculus cDNA clone
            IMAGE:6529663 5', mRNA sequence.
ACCESSION   BU523025
VERSION     BU523025.1  GI:22833475
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14130 row: c column: 07
            High quality sequence stop: 668.
```

FEATURES   Location/Qualifiers
    source   1..906
                 /organism="Mus musculus"
                 /mol_type="mRNA"
                 /strain="FVB/N"
                 /db_xref="taxon:10090"
                 /clone="IMAGE:6529663"
                 /lab_host="DH10B (T1 phage-resistant)"
                 /clone_lib="NCI_CGAP_Co24"
                 /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                 Average insert size 1.6 kb. Constructed by Life
                 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match   39.8%; Score 747.8; DB 13; Length 906;
Best Local Similarity 97.3%; Pred. No.1.9e-76;
Matches 795; Conservative   0; Mismatches 12; Indels 10; Gaps   3;

Qy   52 AGGTAGCAGTGACCCCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTGTAGAGCGA 111
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGTACCAGTGACCCCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTGTAGAGCGA 60

Qy   112 GTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCGAAGAAG 171
         |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   61 GTGTTGCCATGGAACCGATC------AGTATATACACTTCTGATAACTACTCTGAAGAAG 114

Qy   172 TGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATT 231
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   115 TGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATT 174

Qy   232 TCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCA 291
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   175 TCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCA 234

Qy   292 ATGGATTGGTGATCCTGGTCAGTGGTTACCAGCAAGAAGCTAAGGAGCATGACGGACAAGT 351
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   235 ATGGATTGGTGATCCTGGTCAGTGGTTACCAGCAAGAAGCTAAGGAGCATGACGGACAAGT 294

Qy   352 ACCGGCTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCTGGGCAG 411
         |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   295 ACCGGCTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCTGGGCAG 354

Qy   412 TTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCT 471
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 TTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCT 414

Qy   472 ACACTGTCAACCTCACACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACC 531
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 ACACTGTCAACCTCACACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACC 474

Qy   532 TGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAG 591
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 TCGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAG 534

Qy   592 TCTATGTGGGCGGTCTTGGATCCCAGCCCTCCTCCTGACTTCATCTTTGCCG 651
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 TCTATGTGGGCGGTCTTGGATCCCAGCCCTCCTCCTGACTTCATCTTTGCCG 594

Qy   652 ACGTCAGCAGGGGACATCAGTCAGGGGGATGACAGTGACATCTGGACCGGCCTTTACC 711
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   595 ACGTCAGCAGGGGACATCAGTCAGGGGGATGACAGTGACATCTGGACCGGCCTTTACC 654

Qy   712 CCGATAGCCTGTGGATGGTGTGTTCAATTCCAGCATATAATGT-GGTCTCATCCTG 770
         |||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   655 CCGATAGCCTGTGGATGGTGTGTTCAATTCCAGCATATAATGTGGGGTCTCATCCTG 714

Qy   771 CCCGGGCATCGTCATCCTCCTGTTACTGCAATCATCATCTCTAAGCTGTCACACTCCAAG 830
         |||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   715 CCCGGGCATCGTCATCCTCCTGTTACTGCAATCATCATCTCTAAGCTGTCACACTCCAAG 774

Qy   831 GG---CCACCAGAAGCGAAGGCCAGGCCCTCAAGACGCAGAG 864
         || || || ||||||||||||||||||| |||||||||||
Db   CCACCAGAAGCGAAGGCCAGGCCCTCAAGACGACAG 811

RESULT 2
    BX397685
LOCUS   BX397685   1091 bp   mRNA   linear   EST 13-MAY-2003
DEFINITION BX397685 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI041YO13 5-PRIME, mRNA sequence.
ACCESSION   BX397685
VERSION   BX397685.1 GI:30621359
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1091)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI041AH07QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI041AH07QP1.
FEATURES   Location/Qualifiers
    source   1..1091
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="CS0DI041YO13"
                 /tissue_type="PLACENTA COT 25-NORMALIZED"
                 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                 /note="1st strand cDNA was primed with a NotI-oligo(dT)
                 primer. Five prime end enriched, double-strand cDNA was
                 digested with Not I and cloned into the Not I and EcoR V
                 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match   38.4%; Score 720.4; DB 13; Length 1091;
Best Local Similarity 82.1%; Pred. No. 2.1e-73;
Matches 881; Conservative   2; Mismatches 169; Indels 21; Gaps   4;

Qy   38 GCCCGGGCAGGTGCAGGTAGCAGTGACCCTCGAGGCGTTTGGTGCTCCGGTAACCACCA 97
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38 GCTGGTACCGGTCCGGAATTCCCGGGATCGCCGGAGGGCCTGAGTGCTCCAGTAGCACCG 97

Qy   98 CGGCTGTAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACTTCTGATAA 157
         |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   98 CATCTGGAGAACCAGCCGGTTACCATGGA------GGGGATCAGTATATACTTCAGATAA 152

Qy   158 CTACTCTGAAGAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGA 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   153 CTACACCGAGGAAATGGGCTCAGGGACTATGACTCCAGGAGAACCCTGTTTCCGGTGA 212

Qy   218 TGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGAC 277
         ||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   213 AGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTACTTCATCATCTTCCTTAAC 272

Qy   278 TGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGTTCATGGGTTCATGGGTTACAAGGAG 337
         |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   273 TGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGGTTACCAGGAAAG 332

Qy   338 CATGACGGACAAGTACCGGCTGCACCTGTCAGTGGTCGACCTCCTTGTCATCACACT 397
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   333 CATGACGGACAAGTACACAGGCTGCACCTGTCAGTGGTCGACCACCTCCTTGTCATCACGCT 392

cgi-bin/cluster.cgi?seq=CS0DG004AD05QP1&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DG004AD05QP1.

FEATURES   Location/Qualifiers
     source   1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG004YG09"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

ORIGIN

Query Match   37.7%;   Score 707; DB 13;   Length 1201;
Best Local Similarity 75.0%;   Pred. No. 6.7e-72;
Matches 898; Conservative 43; Mismatches 228; Indels 28; Gaps 5;

[DNA sequence alignment between Qy (Query) and Db (Database) sequences, numbered 50/32 through 770, showing aligned nucleotide sequences with match bars]

RESULT 3
BX421287
LOCUS   BX421287   1201 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION BX421287 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG004YG09 5-PRIME, mRNA sequence.
ACCESSION   BX421287
VERSION   BX421287.1 GI:30766068
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/

```
Query Match          37.6%;  Score 705.6;  DB 13;  Length 1201;
Best Local Similarity 82.0%;  Pred. No. 9.7e-72;
Matches  884;  Conservative  7;  Mismatches 161;  Indels  26;  Gaps  6;

Qy     67 TCTGAGGCGTTGGTGCTCCGGTAACCACCACGCGTGTGAGAGCGAGTGTTGCCATGGAAC 126
          || | ||  |||||||||  |||| ||||||||||||||||||||||||||||| |||||
Db     65 TCCGAGGGGCCTGAGTGCTCCAGTAGCCACCACCGCGTGTGAGAGCGAGTGTTACCATGGA- 123

Qy    127 CGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGGAGGTGGGGTCTGAGACT 186
          |           |||||||| ||||||||||||||||||||||||||||||||||||||
Db    124 ----GGGGATCAGTATATACACTTCAGATAACTACTACACCGAGGAAATGGGCTCAGGGGACT 179

Qy    187 ATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAGTTCCATTCAATAGGAGATCTTCC 246
          |||||| |||||||||  |||| ||| |||||||||||| |||| |||||||| |||||||
Db    180 ATGACTCCATGAAGGAACCCTGTTTCCGTTGAAGAAAATGCTAATTTCAATAAAATCTTCA 239

Qy    247 TGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGTGATCC 306
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||| |||||
Db    240 TGCCCACCATCTACTTCCATCATCTTCTTAACTGGCATTGGGCAATGGATTGGTCATCC 299

Qy    307 TGGTCATGGGGTTACCAGAAGGAAGCTTAAGGAGCATGACGGACAAGTTACCGGCTGCACCTGT 366
          |||||||||||| |||||||||  |  ||  ||||||||||||||||||| ||||| |||||||
Db    300 TGGTCATGGGTTACCAGAAGGAAAACTGAGAAGCATGACGGACAAGTTACAGGCTGCACCTGT 359

Qy    367 CAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTCGGCCAGTTGATGGCCATGGCTG 426
          ||||| ||||||||| ||||||||||||||||||||||||||||||||||||||| |||||||
Db    360 CAGTGGCCGACCTCCTCTTTTGTCATCATCACGGTTCCCTTCTCGGGCAGTTGATGGCCGTGGCAA 419

Qy    427 ACTGGTACTTTGGGAAATTTTGTGTAAGGTGTCCATATCATCTACACGTCAACCTCT 486
          ||||||||||||||||| |||| | |||||||| |||||| | ||||||||| |||||
Db    420 ACTGGTACTTTGGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACAGTCAACCTCT 479

Qy    487 ACAGCAGCGTTCTCATCCTCGGCCTTCATCAGCCTGGACCCGGATGCCGTGTGGA 546
          |||||||||| |||||||||||||||||||||||| || ||||||||||||||||
Db    480 ACAGAGTGTCCTCATCCTCGGCCTTCATCAGTCTGGACCCGTACCGTTCCACG 539

Qy    547 CCACCACAGTCAAAGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGGCTCT 606
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CCACCACAGTCAGAGGCCAAGGAGAGCCAAGGAAAGGTGGCTGTATGTTGGCGTCT 599

Qy    607 GGATCCAGCCCTCCTGACTATACCTGACTTCATCATCTTTGCCGACGTGCAGCCAGGGGG 666
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    600 GGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCATCTTCATCTTTGCCAACG------- 647

Qy    667 ACATCAGGGGATGACAGGTACATCTGTACGCCTTACCCCGATGCCGTGTGGA 726
          ||||||||| ||| ||||||||||  || |||||||| | |||||||| |||||
Db    648 ---TCAGTAGGGCAGATGACAGAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGG 704

Qy    727 TGGTGTGTTCAATTCCAGCATATAAATGGTGGTCTCATCCGGCCGGCATCGTCATCC 786
          ||||||||||||||||||||| ||| |||||||||||||||||| ||| ||| |||||
Db    705 TGGTTGTGTTCCAGTTTCAGCACATCATCATCATGGTTGGCCTTATCCTCGTGGTATTGTCATCC 764

Qy    787 TCTCCTGTTACTGCATCATCATCTCCAAGTGCTCCAACTTCCACACTCCAAGGCCCACCAGAGCGCA 846
          | ||||| ||| |||||||||||||| ||||||||||||||||||||| |||| ||||||||||||||
Db    765 TGTCCTGCTATTGCATTATCATCTCCAAGTGTCCACTCTCCAAGGCCTCAACTTCTCCACCAGAAGCGCA 824

Qy    847 AGGCCCTCAAGACGACAGTCATCATCCTCAGCTTTCTTGCTGTCGCTGGCTGCCATATT 906
          |||||||||||||||||  |||||||||| ||| |||| |||||||||||||| |||||
Db    825 AGGCCCTCAAGACCACAGTCATCGATCATCCTCTGGCTTGCTCGCTGTTGGCTGCCTTACT 884

Qy    907 ATGTGGGGATCAGCATCGACTTCTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACT 966
          ||| | | |||||||||||||||| |||||||||||| ||  ||||||||||||||| |
Db    885 ACATTGGGATCAGCATCGACTTCATCCTTCCTGGAAATCATCAAGCAAGGGTGTGAGT 944

Qy    967 TCGAGAGCGCATTGTGCACAAGTGGATCCTCATCACAGGGTCCTTCTTCCACTGTT 1026
          | |||| || |||  | ||| ||||| | ||||:|||||||| ||||||||||||||
Db    945 TTGAGAACACTGTGCACAAGTGGAT-TYCATCACCGAGGCCTAGCTTTCTTCCACTGTT 1003

Qy   1027 GCCTGAACCCATCCCTATGCCTTCCTCGGGCCAAGTTCAAAAGCTCTGCCCAGCATG 1086
          ||   | |  ||||| |||||||||| |||| |   ||  ||  |||||||
Db   1004 GTCTGAA-CCCATCCCTATGCTTTCCTTGGRGCAAATTTAAAACTCT---GCCARCACG 1059

Db    732 GCCTGGTATTGTCATCCTGTCCTGCTATTGCATTATCATCTTCCAAGCTGTCACACTCCAA 791
          |||||| || | |||| | ||  |  ||  |  |||  ||| ||||||||| ||||||||||
Qy    830 GGGCCACCAGAGGCGCAGGCCCTCAAGACGACAGCAGAGCAGTCATCCTCATCCTTCTTTGC 889

Db    792 GGGGCCACCAGAGCGCAGGCGCAGGCCCTCAAGACGACCAGCACCAGTCATCCTCTGGCTTCTCGC 851
          |||||  |||||||||||||: |:||||||||||||||||  ||  ||||||||||:|| |:|
Qy    890 CTGCTGCTGCCATATTATGTGGGGATCAGCAGGATCAGACTGACTCCTTCATCCTTTGGAGTCAT 949

Db    852 CTGTGGCTGCCTTACTGCATCAGCATCGACTCGACCTCCTTCATCCTCCTGGAAATCAT 911
          | |||| |||| : |||||| |:|| ||||| |:  | || : ||  | || : |:||||
Qy    950 CAAGCAAGGATGATGACTCGAGAGCATTGTGCACAAGTCTGCGACAAGAGCCCCT 1009

Db    912 CAAGCAAGGGTGTGAGTTTGAGAACACTGTGCACAGTGGATTYCATCACCGAGGCCCCT 971
          ||||||||| | | |:||| |:|| |::|||||||   |  | ::||||||  |||||||
Qy   1010 CGCCTTCTTCCACTGTTGCCTGAACCCATCCCTCATGCCTTCCTCGGGGGCCAAGTTCAA 1069

Db    972 AGCTTTYTTYCACTGTTGTGAACCCWTCYTCTATG--CTTYCTTGGRGRGCMAATTTWAA 1030
          :|| | ::| ||||||| | | ||| ||  | |||  ||: ||||| | ||:|||:||:||
Qy   1070 AAGCTCTGCCCAGCATCGACTCCAACTCCATGAGCAGAGGCTCCAGCCTCCAGAATCCTTTC 1129

Db   1031 MCTCT---GCCCASACGSCMCTCACCYYTGTRAGCRCRGGGGGCMRCCTCARA----TCYYTC 1083
          :|||      |  |  |    | || ::||| |  | |||   |:||| ||||    ||::||
Qy   1130 CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGGAGTCAGAATCCTCCAGTTTTTCA 1189

Db   1084 CAARGRAAGGCRGGGGGAWTHTYTTTYMKGGTTGGTTWATTTHYCCACYAAAAAAD 1143
          ||| | ||||  ||||| ::|:|: ||:|||| |||| :|||:| |:|| ||||||
Qy   1190 CTCCAGTCAACCCTTATGCAAAGACTTATATATATAAATATATATATGATAAAAGAA 1246

Db   1144 KTAARAT^TTTTTWTRRAAAAHVTTTTTTWTTWWATTTTTDWWAARARGWMAAWKKA 1200
```

RESULT 4

```
BX396365
LOCUS          BX396365  Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION     BX396365 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
               clone CS0DI015YL20 5-PRIME, mRNA sequence.
ACCESSION      BX396365
VERSION        BX396365.1  GI:30609085
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1  (bases 1 to 1201)
  AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE        Full-length cDNA libraries and normalization
  JOURNAL      Unpublished (2001)
COMMENT        Contact: Genoscope
               Genoscope - Centre National de Sequencage
               BP 191 91006 EVRY cedex - France
               Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
               Library was constructed by Life Technologies, a division of
               Invitrogen. This sequence belongs to sequence cluster 1272.r For
               more information about this cluster, see
               http://www.genoscope.cns.fr/
               cgi-bin/cluster.cgi?seq=CS0DI015DFI0QP1&cluster=1272.r. Contact :
               Feng Liang Email: fliang@lifetech.com URL :
               http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
               Faraday Avenue Genoscope sequence ID : CS0DI015DFI0QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI015YL20"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

Qy 1087 CACTCAACTCCATGAGCAGAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGG 1144
Db 1060 CACTCACCTCTGTTAGCAGAGGG3CMAG-CTCAAGATCTCTCMAARGCAAGCGRGGG   1116

LOCUS   BX421511   1201 bp   mRNA   linear   EST 13-MAY-2003
DEFINITION BX421511 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG007YM2 5-PRIME, mRNA sequence.
ACCESSION   BX421511
VERSION   BX421511.1 GI:30651194
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DG007BG06QP1&cluster=1272.r. Contact :
            Feng Liang Email: fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DG007BG06QP1.
FEATURES   Location/Qualifiers
     source   1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG007YM12"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

ORIGIN

Query Match   37.5%;   Score 704.6; DB 13; Length 1201;
Best Local Similarity 81.4%;   Pred. No. 1.3e-71;
Matches 850; Conservative   7; Mismatches 167; Indels 20; Gaps 3;

Qy   31 AGGCGGCGCCGCCGGGGCAGGTGCAGGTAGCAGTGAACCCTTGCGACGGCCGTTGGTGCTCCCGGTA   90
Db   37 ACCGGTCCGGMATTCCCGGGATGGTAGCAAAGTGACGCCGAGGGGCCTGAGTGCTCCCAGTA   96

Qy   91 ACCACCACGGCGTGTAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTT   150
Db   97 GCCACCGCGMTCTGGAGGAACCCAGCCGCGGTTACCATGGA----GGGGATCAGTCAGTATATACACTT 151

Qy 151 CTGATAACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCT   210
Db 152 CAGATAACTACMCCGAGGAAAATGGGCTCAGGGACTCATGACTCCATGAAGGAACCCTGTT   211

Qy 211 TCCGGGATGAAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCT   270
Db 212 TCCGTGAAGAAAATGCTAATTTCAATAAAAMTCTTCCTGCCCACCATCTACTTCATCATCT   271

Qy 271 TCTTGACTGGCATAGTCGGCAATGAGTGGTGATCCGGTCATCCGGTCATCCTGGTTACCAGAAGAGC 330
Db 272 TCTTAACTGGCATTGTGGGCAATGAGTGGTGATCCGGTCATCCGGTCATCCTGGTTACCAGAGAMAC 331

Qy 331 TAAGGAGCATGACGGACAAGTACCGGCTCCTGTCAGTGGCTGACCTCCTCTTTGTCA   390
Db 332 TGAGAAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCA   391

Qy 391 TCACACTTCCCCTTCTGGCAGTTGATGCCATGGCTGACTGACTTTGGAAATTTTGT   450
Db 392 TCACGCTTCCCTTCTGGCAGTTGATGCCGTGGCAAACTGTACTTTGGGAACTTCCTAT   451

Qy 451 GTAAGGCTGTCCATATCATCACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGCCT   510
Db 452 GCMAGGGAGTCCATGTCATCTACACAGTCAACCTCTACAGCMGTGTCCTCATCCTGCCT   511

Qy 511 TCATCAGCCTGGACCGGTACCTGCCCATTGTCCACGGCCACCACAGTCAAAGGCCAAGGA   570
Db 512 TCATCAGTCTGGACCCGCTACCTGGCCATCGTCCACGGCCACCMACAGTCAGGAGGCCAAGGA 571

Qy 571 AACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTTCTGGGGTGGTTCAATTCCAGCACTA 630
Db 572 AGCTGTTGGCTGAAAAGGTGGTCTATGTGGGCTCTGGGGTCGTTTCCAGTTTCAGCACTA 631

Qy 631 TACCTGACTTCATCTTTGCCGACGTCAGCGTCAGGGGGACATCAGTCAGGGGATGACAGGT 690
Db 632 TTCCCGACTTCATCTTTGCCAACGTCAGTGA--------NGGCAGATGACAGAT   677

Qy 691 ACATCTGAGCCGCTTTACCCCGATAGCTGTGGATGGTGGTGTTTCAATTCCAGCATA   750
Db 678 ATATCTGTGACCCGTTCTACCCCAATGACTTGTGGGTGGTTGTTTCCAGTTTCAGCACA   737

Qy 751 TAATGGTGGTCTCATCCTGCCCGGCATCGTCATCCTCTCGTTACTGCATCATCATCT   810
Db 738 TCATGGTTGGCCTTATCATCTGTCATCCTGTCCTCTGTTATTGCATTATCATCT   797

Qy 811 CTAAGTCGTCACACTCCAAGGGCCACCAGAAGCCAAGGCCCTCAAGACGACAGTCATCC   870
Db 798 CCAAGTCGTCACACTCCAAGGGCCACCAGAAGCCAAGGCCCTCAAGACCACAGTCATCC   857

Qy 871 TCATCCTAGCTTCTTCTTGCCTGGCCGCCATATTATGTGGGATCAGCATCGACTCCT   930
Db 858 TCATCCTGGCTTCTTCTTGCCTGGCCGCCTTACTACACATTGGGATCAGCATCGACTCCT 917

Qy 931 TCATCCTTTGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTGCACAAGTGGA   990
Db 918 TCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGCAGAACACTGTGCACAAGTGGA 977

Qy 991 TCTTCCATCACAGAGGCCCTGCCTTCTTCCACGTGTTGCCTGAACCCCATCCTCTATGCT 1050
Db 978 TTTCCATCACCGAGGCCCTAGCTTCTTCCCACGTGTTCTGA--CCCATCCTCTATGCTT   1036

Qy 1051 TCCTCGGGGCCAAGTTCAAAAGCT 1074
Db 1037 TCTTGGAGCAAATTTAAAACTCT   1060

RESULT 6
BX358346
LOCUS   BX358346   1201 bp   mRNA   linear   EST 05-MAY-2003
DEFINITION BX358346 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI036YI7 5-PRIME, mRNA sequence.
ACCESSION   BX358346
VERSION   BX358346.1 GI:30370215
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

```
      Library was constructed by Life Technologies, a division of
      Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
      http://fullength.invitrogen.com/ InVitrogen Corporation 1600
      Faraday Avenue Genoscope sequence ID : CS0DI036CE09QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI036Y17"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match     37.3%;     Score 701;   DB 13;   Length 1201;
Best Local Similarity   82.8%;   Pred. No. 3.2e-71;
Matches 838;   Conservative   4;   Mismatches 150;   Indels 20;   Gaps   3;

Qy   47 GGTGCAGTAGCAGTGACCCCTCGAGGCGTTGGTCGTCCGTAACCACCACGGCTGTAG 106
           |||||||||||| ||||| || |||| ||||| ||| ||||||||| |||||| ||
Db   59 GATGCAGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTTCCAGTAGCCACCGCATCTGGAG 118

Qy  107 AGGGAGTGTTGCCATGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGA 166
           || ||||  ||||||| ||       |||||||||||||| ||||||||||||| ||
Db  119 AACCAGCCGGTTACCATGGA-----GGGGATCAGTATATATACACTTCAGATAACTACACCGA 173

Qy  167 AGAAGTGGGGTCGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGT 226
           ||||||||||| | ||||||||||| |||     ||||||||||| ||||||||||
Db  174 GGAAATGGGGCTCAGGGGACTATGACTCACTGTTCCGTGAAGAAAATGC 233

Qy  227 CCATTCAATAGAATCTTCCTGCCCACCATCACTTCATCATCTTCTGACTGGCATAGT 286
           |||| ||||  ||||||||||||||||||| ||| |||||||||  |||||||| ||
Db  234 TAATTTCAATAAAATCTTCCTGCCCACCATCCATCACTTCCATCATCTTCTTAACTGGCCATTGT 293

Qy  287 CGGCAATGGCATTGGTGATCCTGGTCATGGGTTACCAGGAGCTAAGGAGCATGACGGA 346
           |||||||| ||||||| ||||||||||||||||||||||||   ||| ||||||||||
Db  294 GGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGGAGAAACTGAAGAAGCATGACGGA 353

Qy  347 CAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTG 406
           |||||||  |||||||||||||||| | ||||||||||||||||||| || ||||||||
Db  354 CAAGTACAGGCTGCACCTGTCAGTGCCGACCTCCTCTTGTCATCACGCTTCCCTTCTG 413

Qy  407 GGCAGTTGATGCCATGGCGTCTGACTGGTACTTTGGGAAAATTTTGTGTAAGGCTGTCCATAT 466
           |||||||||||| || |  ||  ||||||||||||||||| ||   |  ||||| |||||| |
Db  414 GGCAGTTGATGCCGTCGGTGGCAAACTGGTACTTTGGGAAACTTCCTATGCAAGGCAGTCCATGT 473

Qy  467 CATCTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCG 526
           |||||||| |||||||||||||||||| | ||| |||||||||||||||||| ||||||||
Db  474 CATCTACACAGTCAACCTCTACAGCAGTGTCCTTCATCCTGGCCTTCATCAGTCTGGACCG 533

Qy  527 GTACCTGGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGAAAACTGCGTCTGAAAA 586
           ||||||||||||||||||||||||||||||||| || ||||||| || | |||||||||
Db  534 CTACCTGGCCATTGTCCACGCCACCAACAGTCAAGAGGCCAAGAAGCTGTTGGCTGAAAA 593

Qy  587 GGCAGTCTATGTGGGCGTCTGGATCCCCAGCCCTCCTGACTATACCTGACTTCATCTT 646
           || | ||||||| ||||||||||||||| | ||| |  ||||| ||  |||||||||
Db  594 GGTGTCTATGTTGGCGTCTGGATCCCCGTCCCCTCGCTGCTGACTATTCCCGACTTCATCTT 653

Qy  647 TGCCAGCGTCAGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGGCCT 706
           |||||  |||||  | |         ||| |||||||||||| ||| | |||||||||| |
Db  654 TGCCAACGTCAGTGA----------NGGCAGATGACAGATATATTCTGTGACCGGCTT 699

Qy  707 TTACCCGGATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAAATGGTGGGTCTCAT 766
           ||| || |||| ||||||| ||||||||| ||  ||||  |  ||||| |||| ||  ||
Db  700 CTACCCCAATGACTTGTGGGTTGTGGTGTGTTCAGTTTCAGTGCACATCATGGTTGGCCCTTAT 759

Qy  767 CCTGCCCGGCCATCGTCATCCTCCTCGTTACTGCATCATCATCCTCTAAGCTGTCACTC 826
           ||||||  | | || |||||||||||||||  | ||||| ||||| | |||||||| ||
Db  760 CCTGCCTGCTGTATTGTCATCCTCCTCGTTCGTATTGCATTATCATCTCCAAGCTGTCACACTC 819

Qy  827 CAAGGGCCACCAAGAAGCGGCAAGGCCCTCAAGAGACGAGACAGTCATCCTCATCCTCCTAGCTTCTT 886
           ||||||||||| |||||||||||||||||||||  || |||  |||| ||||  |||  ||| ||||
Db  820 CAAGGGCCACCAGAAGCGGCAAGGCCCTCAAGACCCAGTCATCCTCATCCTGGCTTTCTT 879

Qy  887 TGCCTGCTGCTGCCATATTATGTGGGATGCAGATCAGCATCAGCGACTCCTTCATCCTTTTGGGAGT 946
           ||||||  ||||||| |  |  | |||| ||||||||||||||||||||||||||||||    |||| |
Db  880 CGCCTGTTGCTGCCTTACTACATTGGGATCAGCATCAGCATCAGCGACTCCTTCATCCTCCTCCTGGAAAT 939

Qy  947 CATCAAGCAAGGATGTGACTTCGAGAGACATTGTGCACAAGTGATCTCCATCACAGAGGC 1006
           |||| |  |||| ||||| ||| ||  || |||||||| |||||||| ||||| ||||| ||
Db  940 CATCACCAAGGGGTGAGTTTGAGGTTGGAACTGTGTGCACAAGTGGATTTCCATCACCGAGGC 999

Qy 1007 CCTGGCCTTCTTCCACTGTGCCGAACCCATCCTCTATGCCTTCCTCGGG 1058
           ||| || |||||||||||| |  ||| ||||||||||||      | ||| |
Db 1000 CCTAGCTTTCTTCCCACTGTGTYTGAA-CCCATCCTCTATTSTTTCYTGGGRG 1050


RESULT 7
BX353253
LOCUS       BX353253               1201 bp    mRNA       linear   EST 02-MAY-2003
DEFINITION  BX353253 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC003YG03 5-PRIME, mRNA sequence.
ACCESSION   BX353253
VERSION     BX353253.1   GI:30347726
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DC003AD02QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fullength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DC003AD02QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC003YG03"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match     36.5%;   Score 685;   DB 13;   Length 1201;
Best Local Similarity   80.4%;   Pred. No. 2.2e-69;
Matches 863;   Conservative   6;   Mismatches 182;   Indels 23;   Gaps   5;

Qy   45 CAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGGTGTCTCCGGTAACCACCACGGCTGT 104
           | |||||| ||||||| ||| ||  ||||| ||||| ||| |  ||||||||||||||
Db   60 CGGGATCAGTTAGCAAAGTGACGCCGAGGGGCTGAGTGCTCCAGTAGCCACCGCATCTGG 119

Qy  105 AGAGCGAGTGTTGCCATGGACGACCGATCAGTGTGAGTATATCACTTCTGATAACTACTCT 164
           ||| | || ||||||| || | |        ||||||||||| |||||||||||||||
Db  120 AGAACCAGCGGTTACCATGGA----GGGGATCAGTATATATACACTTCAGATAACTACACC 174

Qy  165 GAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAAC 224
```

ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1051)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12720 row: e column: 12
            High quality sequence start: 753.
            High quality sequence stop: 753.
FEATURES   Location/Qualifiers
     source   1..1051
                     /organism"Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5726963"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_124"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV
                     (destroyed); Site 2: NotI; RNA source male hippocampus,
                     age 27. Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.4 kb, insert size range 0.9-4 kb. Library is
                     normalized and enriched for full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 012."

ORIGIN

Query Match   36.5%; Score 684.8; DB 12; Length 1051;
Best Local Similarity   81.6%; Pred. No. 2.5e-69;
Matches   873; Conservative   0; Mismatches 172; Indels   25; Gaps   6;

[sequence alignment data]

```
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="CS0DH005YK10"
                    /tissue_type="T CELLS (JURKAT CELL LINE)"
                    /cell_line="JURKAT CELL LINE"
                    /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
                    /note="vector: pCMVSPORT 6; 1st strand cDNA was primed
                    with a NotI-oligo(dT) primer. Five prime end enriched,
                    double-strand cDNA was digested with Not I and cloned into
                    the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                    Library was not normalized."

ORIGIN

        Query Match      35.2%;     Score 660.8;   DB 13;   Length 1201;
        Best Local Similarity   82.3%;    Pred. No. 1.2e-66;
        Matches 815;   Conservative   0;   Mismatches 152;   Indels   23;   Gaps    4;

Qy     55  TAGCAGTGACCCTGAGGCGTTGGTGTCCGGTAACCACCACGGCTGTAGAGCGAGTG  114
           |||||        ||||| |||||   || |||  || |||| ||| ||||  |||
Db     67  TAGCAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGAGAACCAGCG  126

Qy    115  TTGCCATGGAACCGATCAGTCAGTGAGTATATAGACCTCTGATAACTACTCTGAAGAGAGTGG  174
           ||| |||||||      |||||||||||  |||||   |||||     ||||| ||| ||||
Db    127  GTTACCATGGA------GGGGATCAGTATATACACTTCAGTAATACACCGAGAAATGG  181

Qy    175  GGTCTCGGAGACTATGATCCAACAAGAGAACCCTGCTTCCGGGATGAAAAGTCCATTTCA  234
           |||| ||||||||||||  ||  |||| ||||| ||| ||| ||||| |||||| ||||
Db    182  GCTCAGGGGACTATGACTCCATGAGAAGGAACCCCTGTTTCCGTGAAGAAAATGCTATTTCA  241

Qy    235  ATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCATCATCTTCTTAACTGGCATGGCATCGGCAATG  294
           ||    |||||||||||||||||||||||||||||||||||||||||||||||||| || ||| ||||
Db    242  ATAAAAATCTTCCTGCCCACCATCTACTTCATCATCTTCATCATCTTAACTGGCATTGTGGGCAATG  301

Qy    295  GATTGGTGATCCTGGTTACCAGAAGAGCTAAGGAGCATGACGGACAAGTACC  354
           ||||||||||||||||||||||||||||||  ||||||| |||||||||| ||
Db    302  GATTGGTCATCCTGGTTACCAGAAGAGCATGACGGACAAGTACA  361

Qy    355  GGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCAACACTCCCCTTCTGGGCAGTTG  414
           ||||||||||||||||||| ||||||| |||||||||  |||||| |||||||||||||
Db    362  GGCTGCACCTGTCAGTGGCCGACCTCCTCTTCTGTCATCAACGCTTCCCTTCTGGGCAGTTG  421

Qy    415  ATGCCATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGCAAGGCAGTCCATCATCACA  474
           |||| ||||||| ||||| || |||||||||||||  ||||||||||||||| ||||||||||
Db    422  ATGCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGCAGTGCCATGTCATCATCACA  481

Qy    475  CTGTCAACCTCACAGCAGCGGTTCTCATCCTCGGCCTTCATCAGCCTGGACCGGTACCTCG  534
           |||| |||||||||||||||||| |||||||| ||||||||||||| |||||||||||||
Db    482  CAGTCAACCTCACAGCAGCGGTTCTCATCCTCGGCCTTCATCAGCCTGTCGACCGTACCTGG  541

Qy    535  CCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGAAACTCGTGGCTGAAAAGGCAGTC  593
           ||||| ||||||||||||||||| |||||||||||||||||||||||||||||| ||
Db    542  CCATCGTCCAGCCACCAACAGTCAGAGGCCAAGGAAACTCGTGGCTGAAAAGGTGGTC  601

Qy    594  TATGTGGGCGTCTGGATCCCAGCCCTCCTCTGCTGACTATACCTGACTTCATCTTGCCGAC  653
           ||||| |||||||||||||| |||||||| ||||||||||| || ||||||||||||||
Db    602  TATGTTGGCGTCTGGATCCCGCCCTCCTGCTGCTGACTATTCCCGACTTCATCTTGCCAAC  661

Qy    654  GTCAGCCAGGGGACACATCAGTCAGGGGATGACAGAGTACATCTGTGACCGGCCTTTACCCC  713
           |          |||| |||| |||| ||||| |||||||| |||||||||||| || ||||||
Db    662  G---------TCAGTGAGGCAGCGATGACGATGACAGCAGATAATCTGTGAAAACAAAGGGGTCCAGCCTCA  706

Qy    714  GATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAAATGGTGGGTCCATCCTGCCC  773
           ||||| |||||| |||||||||||||| ||||| ||||| |||||| |||||||||||
Db    707  ATGACTTGTGGCGTGGTGGTGTGTTGTTCAGTTCAGTTGCACATCATGGTTGGCCTTATCCTGCCT  766

Qy    774  GGCATGTGCATCCTCTCTGTTACTGTGATCATCATCCTCTAAGCTGCTCCTAGCTGACTGC  833
           || |||||||||||| ||||| |||| ||||||||||||| | |||||| |||||||||||
Db    767  GGTATTGTCATCCTGTCCTGCTATTGCATTATCATCATCTCCTAGCTTTGCTTTGCTGC  826

Qy    834  CACCAGAAGCGCAAGGCCCTCAAGACGACACTCGAAGACGACAGTCAGTTCAGTGCT  894
           |||||||||||||||||||||||||| || || ||||| || ||||| ||||||||
Db    827  ACCAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCGACCATCATCCTCTGGCTTTCGCCTGTT  886

Qy    894  TGGCTGCCATATTATGTGGGGATCAGCATGCGACTCCTTCATCTTTGGGAGTCATCAAG  953
           |||| |||||||| || |||| |||||||||| | ||||| ||||||| ||||||||||
Db    821  TGGCTGCCTTACTACATTGGGGATCGACATCCTCTGGCTTTCGCCTGTT  880

Qy    954  CAA-GGATGTGACTTCGAGAGACATTGTGCACAAGTGGATCCTCCATCACAGA-GGCCCTCG  1011
           ||| || |||||| ||||||| |||||| ||| || |||  || ||||| |||  ||||||||
Db    881  CAAGGGGGTGAGTTTGAGAACACGGGTGCACAAGTGGGATTTTCATCACCGAGGGCCCTAG  940

Qy   1012  CCTTCTTCCACTGTTGCTGCCAGCCCATGAATCCTCTATGCTT--CCTCGGGGCCAAGTTCAA  1069
           |||||||||| ||| ||||| ||  ||| |||||||||||||    |||  |||||||||||
Db    941  CTTTCTTCACTGTTTGGCTGAACACGGTCCCATCCTCTATGCTTCCTGGAGCCCCAATTTAA  1000

Qy   1070  AAGCTCTGCCCAGCATGCACTCAACTCCATG-AGCAGAGGCTCCAGCCTC  1118
           ||||| |||||||| ||||   ||||| ||| |||||||||||||||||
Db   1001  AACCTCTGCCCAGCAGCGACCACTCCACTGTCGTGAAACAAAGGGGTCCAGCCTC  1050
```

RESULT 9
BX462627
LOCUS   BX462627   1201 bp   mRNA   linear   EST 22-MAY-2003
DEFINITION BX462627 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
            clone CS0DH005YK10 5-PRIME, mRNA sequence.
ACCESSION   BX462627
VERSION   BX462627.1 GI:31029445
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DH005BF05QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/. Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DH005BF05QP1.
FEATURES   Location/Qualifiers
     source 1..1201
             /organism="Homo sapiens"

```
Qy   219 GAAAACGTCCATTCAATGGATCTTCTGCCACCATCTCACTTCATCATCTTCTGACT 278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   236 GAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTCACTTCATCATCTTCTAACT 295

Qy   279 GGCATAGTGCGGCAATGATTGGTGATCTGGTTCATGGTTACCAGAGAAGTAAGAGAGC 338
             |||||||||||| ||||||||||  |||||||||||||||||||||||||||| |
Db   296 GGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGTTACCAGAAGAAACTGAGAAGC 355

Qy   339 ATGACGGACAAGTACCGGCTGACCTGGCTGACCTCCTCCTTCTTGTCATCACACTC 398
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   356 ATGACGGACAAGTACCAGGCTGCCTGGCTGACCTCCTCCTTCTTGTCATCACGGTT 415

Qy   399 CCCTTCTCGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGAATTTTGTGTAAGGCT 458
             ||||||||||||||||||||||| |||||||||||||||||||||||
Db   416 CCCTTCTCGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCA 475

Qy   459 GTCCATATCATTCACTGTCAACCTCTACACGACAGCTTCTCATCCTGGCCTTCATCAGC 518
             |||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   476 GTCCATGTCATCTCACACAGTCAACCTCTACACGACAGTGTCCTTCATCCTGGCCTTCATCAGT 535

Qy   519 CTGGACCGGTACCTGCCATTGTCCAGGCACCACAACAGTCAAAGGCCAAGGAAACTGCTG 578
             ||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   536 CTGGACCGGTACCTGGCCATCGTCACGGCACCACAACAGTCAGAGCCCAAGGAAGCGTTG 595

Qy   579 GCTGAAAAGGCAGTCTATGTGGCGTCTGGATCCCAGCCTCCTGACTATACCTGAC 638
             |||||||||||||||||||||||||||||||||||||||||||||||
Db   596 GCTGAAAAGGTGGTTCTATGTTGGCGTCGTCGGTTCCAGTTCAGCACACATCATGGTT 655

Qy   639 TTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGGGATGACAGTACAGTACATCTGT 698
             ||||||||||||||||                    |||||||||||||||||||||||
Db   656 TTCATCTTTGCCAACG------------TCAGTRAGGCAGATGACAGATATATCTGT 700

Qy   699 GACCGCCTTACCCGCATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTG 758
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   701 GACCGCCTTCTACCCCAATGACTTGGGTGGTGTGTTCCAGTTCCAGTTCAGCACACATCATGTT 760

Qy   759 GGTCTCATCCTCGCCCGGACGTCATCCTCCTCGTTACTGCATCATCATCTCTAAGCTG 818
             |||||||||||||||| |||||||||||||||| |||||||| |||||||||
Db   761 GGCCTTATCCTCGCCTGCTGGTCTCGTCCTGCTATTGCATTATCATCCTCCAAGCTG 820

Qy   819 TCACACTCCAAGGGCCACCAGAAGGCCAAGGCCCTCAAGACCAGCAGTCATCCTCATCCT- 877
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   821 TCACACTCCAAGGGCCACCAGAAGGCCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTG 880

Qy   878 AGCTTTCTTTGCGCTGCTCGGCTGCCATATTATGTGGGGATCAGCATGCATCAGCTCCTTCATCCT 937
             ||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   881 GGCTTTCTTCGCCTGTTGGCTGCCTTACTACATTGGGGATCAGCATGCATCAGCTCCTTCATCCT 940

Qy   938 TTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGACGATTCGACAAGTGGATCTCCAT 997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   941 CCT-GGAAAYATCAAGCAAGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGAT-TCCAT 998

Qy   998 CACAGAGGCCCTGCCTGTTCTTCCACTGTTGCCTGAACCCCATCCTCT 1044
             ||||||||||||| |||||||| || ||||||||||||||||||||
Db   999 CACCGGAGGCC--TAGCTTTCTCYCMCTGTTGTCTGAACCCATCCTCTAT 1043
```

```
Qy   895 GGCTGCCATTATGTGGGGATCAGCATGCGACTGACTCCTTCATCCTTTGGGAGTCATCAAGC 954
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   887 GGCTGCCTTACTACATTGGGGATCAGCATGCATCGACTGACTCCTTCATCCTTGGAAAATCATCAAGC 946

Qy   955 AAGGATGTGACTTCGAGAGCATTCGAGACACTGTGCACAAGTGGATCTCCATCCTGAACCCCATCCTCT 1014
             |||| ||||||||||||||||||||| |||||||||||||| |||||||||
Db   947 AAGGGTGTGAGTTTGAGAACACTGTGCACAAGTGGAT-TCATCACCGGAGGCC--TAGCT 1003

Qy   1015 TCTTCCACTGTTGCCTGAACCCCATCCTCT 1044
              |||||||||||||||||||||||||||||
Db   1004 TTCTTCACTGTTGTCTGAACCCCATCCTCTAT 1033
```

RESULT 10
```
BX418530
LOCUS       BX418530              1201 bp    mRNA    linear   EST 13-MAY-2003
DEFINITION  BX418530 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF006YJ16 5-PRIME, mRNA sequence.
ACCESSION   BX418530
VERSION     BX418530.1  GI:30658606
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr  Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DF006DE08QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF006DE08QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF006YJ16"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN
```

```
Query Match    34.5%;  Score 647.4;  DB 13;  Length 1201;
Best Local Similarity  81.7%;  Pred. No. 4.1e-65;
Matches 823;  Conservative   4;  Mismatches 155;  Indels  25;  Gaps   6;

Qy   39 CCCGGGCAGGTGCAGGTAGCAGTGACCCCTGAGGCGTTGGTGTCCGGTAACCACCACAC 98
            ||||||||||||| ||||||||| |||||||||||||||||||||||||||||
Db   61 CCCGGGGATGCGAGCACAGGTAGCAGTGACCGGAGGCCTGAGGGCTGCTCCAGTAGCCACCGC 120

Qy   99 GGCTGGTAGAGGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTAATAGACTTCTGATAAC 158
            |||||||||||||||||||||||||||||||||||          ||||||||||
Db   121 ATCTGAGAACCAGCCGGTTACCATGGAGGGGGA-----TCAGTATATACACTTCAGATAAC 175

Qy   159 TACTTCTGAAGAAGTGGGGTCTCGGAGGACTATGACTCCAACAAGGAACCCTGCTTCCCGGGAT 218
            |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   176 TACACCGAGGAAATGGGCTCAGGGGACTCATGAYTCCATGGGAACCCTGTTTCCGTGAA 235
```

RESULT 11
```
BG174412
LOCUS       BG174412                958 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602334232F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457694 5',
            mRNA sequence.
ACCESSION   BG174412
VERSION     BG174412.1  GI:12681115
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 958)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10254  row: o  column: 07
            High quality sequence stop: 628.

FEATURES             Location/Qualifiers
     source          1..958
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4457694"
                     /tissue_type="tumor; biopsy sample"
                     /dev_stage="10 months; virgin"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam1"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match       33.6%;    Score 631.4;  DB 12;  Length 958;
Best Local Similarity   96.2%;   Pred. No. 3.2e-63;
Matches 690;  Conservative   0;  Mismatches  21;  Indels   6;  Gaps   4;

Qy    54 GTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGCGAGT 113
         ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     1 GTAGCAGTGACCCCTCTGAGGCGTTGGTGTCCGGTAACCACCACGGCTGTAGAGCGAGT 60

Qy   114 GTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTG 173
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTG 120

Qy   174 GGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTC 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTC 180

Qy   234 AATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTTGACTGGCATGGTCGGCAAT 293
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATAGGATCTTCCTGCCCCACCATCACTTCATCATCTTCTTGACTGGCATGGTCGGCAAT 240

Qy   294 GGATTGGTGATCCGGTCATGGGGTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTAC 353
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGATTGGTGATCCGGTCATGGGGTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTAC 300

Qy   354 CGGCTGCACCCGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCTGGGCAGTT 413
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CGGCTGCACCCGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCTGGGCAGTT 360

Qy   414 GATGCCATGGCTGACTGGTACTTTGTGGAATTTTTGTGTAAGGCGTGTCCAATATCATCTAC 473
         |||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   361 GATGCCATGGCTGACTGGTACTTTGGGAAA-TTTTGTGTAAGGCGTGTCCATATCATCTAC 419

Qy   474 ACTGTCAACCTCTACAGCAGCGTTCTTCATCCTCGGACCCGTACCTC 533
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   420 ACTGTCAACCTCTACAGCAGCGTTCTTCATCCTCGGACCCGTACCTC 479

Qy   534 GCCATTGTCCAGGCCACCAAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTC 593
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   480 GCCATTGTCCAGGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTC 539

Qy   594 TATGTGGGGCTCTGATCCCAGCCCTCCTCCAGCCCTCCGACTATACCTGACTTGCCGAC 653
         |||||||| |||||||||||||||||||||| |||||||||||||||||||||||||||
Db   540 TATGTGGGGCTCTGATCCCAGCCCTCCTCCGA-TATACCTGACTTCATCTTTGCCGAC 598

Qy   654 GTCAGCCAGGGGGAGCATCAGCGGGGATGACAGCAGGTACATCGTGGACCGGCCTTACCCCGA 713
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   599 GTCAGCCAGGGGGACATCAGCAGGGGATGACGAGGTACTTCTGTGACGACGGCTTACCCGA 658

Qy   714 GATAGCCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTG 770
         ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   659 TA--GCTGTGGATGGGGGCG-TTCAACTCCAGCATATTAAGGGGGTCCATCCTGCCG 711


RESULT 12
LOCUS       BX462182                1201 bp    mRNA            linear  EST 22-MAY-2003
DEFINITION  BX462182 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG003Y019 5-PRIME, mRNA sequence.
ACCESSION   BX462182
VERSION     BX462182.1  GI:31031389
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DG003AH10QP1&cluster=1272.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DG003AH10QP1.

FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG003Y019"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

ORIGIN

Query Match       33.4%;    Score 626.6;  DB 13;  Length 1201;
Best Local Similarity   78.0%;   Pred. No. 9.7e-63;
Matches 785;  Conservative  29;  Mismatches 166;  Indels  27;  Gaps   5;

Qy    45 CAGGTGCAGGTAGCAGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGT 104
         || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    50 CGGGATCAGGTAGCAAAGTGACGCCCAGGGCCTGAGTGCTCCGAGTGTCCCACCGCATCTGG 109

Qy   105 AGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCT 164
         |||||||||||||||||||||||||||||||||||| ||||| |||||||||||||||||
Db   110 AGAACCAGCGCGTTACCATGGA----GGGGATCGATATATACACTTCAGATAACTACCACC 164

Qy   165 GAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAAC 224
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   165 GAGGAAATGGGCTCAGGGGACTATGGCTCCCATGAGCCATGGAACCCTGTTTTCCGTGAGAAAT 224

Qy   225 GTCCATTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTTGACTGGCGTGT 284
         ||||||||||||| |||| ||||||||||||||||||||| ||||| |||||||||||||
Db   225 GCTAATTTCAATAAAATCTTCCTGCCCACCATCACTTCCATCATCTTTAACTGGCATT 284

Qy   285 GTCGGCAATGGATGGTGGTCATCCGGTCATGGGGTGATGACAGCAGGTTACCAGAAGAAGCTAAGGAGCATGACG 344
```

```
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:388497
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 497.

FEATURES    Location/Qualifiers
     source         1..793
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /strain="C57BL/6J"
                    /db_xref="taxon:10090"
                    /clone="IMAGE:636505"
                    /sex="male"
                    /tissue_type="lymph node"
                    /dev_stage="4 weeks"
                    /lab_host="DH10B"
                    /clone_lib="Soares mouse lymph node NbMLN"
                    /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)
                    with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                    1st strand cDNA was primed with a Not I - oligo(dT) primer
                    [5'
                    TTTACCAATCTGAAGTGGGAGCGGCCGCCGCCATTACTTTTTTTTTTTTTTTTTTTTTTTTT
                    3'], double-stranded cDNA was ligated to Eco RI adaptors
                    (Pharmacia), digested with Not I and cloned into the Not I
                    and Eco RI sites of the modified pT7T3 vector. RNA
                    provided by Dr. Bertrand Jordan. Library constructed and
                    normalized by Bento Soares and M.Fatima Bonaldo."

ORIGIN
       Query Match          33.1%;   Score 621.6;   DB 9;   Length 793;
       Best Local Similarity  92.4%;  Pred. No. 4.9e-62;
       Matches  715;  Conservative  0;  Mismatches  46;  Indels  13;  Gaps  6;

Qy   633 CCTGACTTCACTCTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGGGATGACAGGTAC 692
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CCTGACTTCACTCTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGGGATGACAGGTAC 60

Qy   693 ATCTGTGACCGCCTTACCCGATAGCCTGTGGATGTGGTGTTTCAATTCCAGCATATA 752
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATCTGTGACCGCCTTACCCGATAGCCTGTGGATGTGGTGTTTCAATTCCAGCATATA 120

Qy   753 ATGGTGGTCTCATCCTGCCCGGCATCGTCATCCTCTCCTGTTACTGCTCATCATCCT 812
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCTCCTGTTACTGCTCATCATCTCT 180

Qy   813 AAGCTGTCACACTCCAAGGGCCACCAGAAGCGCCAAGGGCCCTCAAGACGACAGTCATCCTC 872
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AAGCTGTCACACTCCAAGGGCCACCAGAAGCGCCAAGGGCCCTCAAGACGACAGT-ATCCTC 239

Qy   873 ATCCTAGCTTTCTTTGCCTGCTGCTGCTGCCATATATGTGGGGATCAGCATCAGCGACTCCTTC 932
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240 ATCCTAGCTTTCTTTGCCTGCTGCTGCCATATTATGTGGGGATCAGCATCAGCGACTCCTTC 299

Qy   933 ATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTGAGAGACATTGTGCACAAGTGGATC 992
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   300 ATCCTTTGGGGGTCATCAAGCAAGGATGTGACTTCGAGAGACATCGTGCACAAGTGGATC 359

Qy   993 TCCATCACAGAGGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTTC 1052
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 TCCATCACAGAGGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTTC 419

Qy  1053 CTCGGGGCCAAGTCAAGTCAAAAGTCTGCCCAGCATGCACTCAACTCAACTCCATGCAGAGGCTCC 1112
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   420 CTCGGGGCCAAGTCAAGTCAAAAGTCTGCCCAGCATGCACTCAACTCCATGCAGAGGCTCC 479

Qy  1113 AGCCTCAAGATCCTTCCAAAGGAAAGCGGGGGTGGACACTCTTCCGTCTCCACGGAGTCA 1172
```

```
Db   285 GTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGGAAACTGAGAAGCATGACG 344
Qy       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         GACAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCTTC 404

Db   345 GACAGTACCGGCTGCACCTGTCAGTGGCCGCNCCTCCTTCTTTGTCATCACGCTTCCCTTC 404
Qy       |||||||| |||| ||||||||||||||| |||||||||| |||||||||| |||||

Db   405 TGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCAT 464
Qy       TGGGCAGTTGATGCCATGGCTGACTTGGGAAACTTGGTACTTTGGGAAACTGGTACTCCTATGCAAGGCAGTCCAT 464

Db   465 ATCATCTACTACTGTCAACCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGAC 524
Qy       GTCATCTACACACANTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGAC 524

Db   525 CGGTACCTCGGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAA 584
Qy       CGCTACCTGGCCATCGTCCACGCCACGTCCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAA 584

Db   585 AAGGCAGTCTATGTGGGCGTCTGGATCACCAGCCCTCCTCGACTATACCTGACTTCATC 644
Qy       AAGGKGGKCTATGTTGGCGTCTGGATCACCAGCCCTCCTCGACTATACCTGACTTCATC 644

Db   645 TTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGC 704
Qy       TTTGCCAACG--------TCAGTAAGGCAGATGACAGATATATCTGTGAMCGC 689

Db   705 CTTACCCGATAGCCTGTGGATGTGGTGTTTCAATTCCAGCATATAATGGTGGGTCTC 764
Qy       TTCTACCCCAATGACTTGTGGGKGGKGTGTTGTTCCAGTTTMAGCAMATMATGGTGGCCCT 749

Db   765 ATCCTGCCCGGCATCGTCATCCTCTCCTGTTACTGCTCATCATCTCTAAGCTGTCACAC 824
Qy       ATCATMAAGCAAGGAGTGAGTTTRARA-MACTGTKCAMAATGGRTT-----TCAYAMCG 982

Db   825 TCCAAGGGCCACCAGAAGCGCCAAGGGCCCTCAAGACGACAGTCATCCTCAGCTTTC 884
Qy       GCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTT 1051

Db   810 TCCAAGGGSCACCAGAAGCGCCAAGGGCCCTCAAGACGACAGTCATCCTCATCCTGGCTTTC 869
Qy       AGGCCYARCTTTYTTYTTCMMTKTKTTTKACCCAHCYCTWTGTTTYTTT 1029

Qy   885 TTTGCCTGCTGCTGCCAATATTATGTGGGGATCAGCATCGACTCCTTCATCCTTTGGGA 944
Db   870 TTTGCCTGTTGGCTGCCTTACTACATTGGGATMA-SATCCACTCCTTCATCCTCCTGAA 928

Qy   945 GTCATCAAGCAAGGATGTGACTTCGAGAGACATTGTGCACAAGTGGATCTCCATCACAGAG 1004
Db   929 ATCATMAAGCAAGGAGTGAGTTTRARA-MACTGTKCAMAATGGRTT-----TCAYAMCG 982

Qy  1005 GCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATCCTCTATGCCTT 1051
Db   983 AGGCCYARCTTTYTTYTTCMMTKTKTTTKACCCAHCYCTWTGTTTYTTT 1029

RESULT 13
AA182270
LOCUS       AA182270     793 bp    mRNA        linear   EST 06-JAN-1997
DEFINITION  mt83f01.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
            IMAGE:636505 5' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
            RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION   AA182270
VERSION     AA182270.1  GI:1765810
KEYWORDS    EST
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 793)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE       The WashU-HHMI Mouse EST Project
JOURNAL     Unpublished (1996)
```

characteristic muscle proteins. "

ORIGIN

Query Match   33.1%;   Score 621.2; DB 9;   Length 639;
Best Local Similarity 99.5%;   Pred. No. 6.3e-62;
Matches 623;   Conservative   0; Mismatches   3;   Indels   0; Gaps   0;

Qy   688 GGTACATCTGTGACCGGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAATTCCAGC 747
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   14 GGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAATTCCAGC   73

Qy   748 ATATAATGGTGGGTCTCATCGTGCCGGACATCGTCATCCTCCTGTTACTGCATCATCA   807
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74 ATATAATGGTGGGTCTCATCGTGCCGGACATCGTCATCCTCCTGTTACTGCATCATCA   133

Qy   808 TCTCTAAGCTGTCACACTCCAAGAGCGCCACCAGAAGCGCAAGCGCCCTCAAGACGCAGCAGTCA 867
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 TCTCTAAGCTGTCACACTCCAAGGGCGCCACCAGAAGCGCAAGCGCCCTCAAGACGCAGCAGTCA 193

Qy   868 TCCTCATCCTCAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGCATCGACT 927
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194 TCCTCATCCTCAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGCATCGACT 253

Qy   928 CCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGACAGCATTGTGCACAAGT 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 CCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGACAGCATTGTGCACAAGT 313

Qy   988 GGATCTCCATCACAGAGGCCCTGCGCTTCTCCACTGTTGCCTGACCCCATCCTCTTATG 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314 GGATCTCCATCACAGAGGCCCTCGCGCTTCTCCACTGTTGCCTGACCCCATCCTCTTATG 373

Qy   1048 CCTTCCTCGGGGCCAAGTCAAAGTCTCGCCAGCAGCACTCAACTCAACTCCATGAGCAGAG 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   374 CCTTCCTCGGGGCCAAGTTCAAAGTCTCGCCAGCATGCACTCAACTCAACTCCATGAGCAGAG 433

Qy   1108 GCTCCAGCCTCAAGATCCTTTCCAAAGGAAGCGGGGTGGACACTCTTCCGTCTCCACGG 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   434 GCTCCAGCGTCGAAGATCCTTTCCAAAGGAAGCGGGGTGGACACTCTTCCGTCTCCACGG 493

Qy   1168 AGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATATAATAATAT 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   494 AGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATATAATAATAT 553

Qy   1228 ATATATATGATAAAGAAACTTTTTATGTTACACATTTCCAGATATAAGAGACTGACC 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   554 ATATATATGATAAAGAAACTTTTTATGTTACACATTTCCAGATATAAGAGACTGACC 613

Qy   1288 AGTCTTGTACAGTTTTTTTTTTTTTTT 1313
              ||||||||||||||||||||||||
Db   614 AGTCTTGTACAGTTTTTTTTTTTTTTTT 639

Db   480 AG-CTCAGAGATCCTTT-CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCACGGAGTCA 537
              || |||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Qy   1173 GAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATATA---ATATATA 1228
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   538 GAATCCTCCAGTTTCACTCCAGCTAACTTATGCAAAGACATATATAATAATATAATATATA 597

Qy   1229 TATATAATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAGACTGACCA 1288
              ||||||| ||| |||| ||||||||||||||||||||||||||||||||||||||||
Db   598 TATATAATGATAGAGAAGCTTTTTATGTTACACATTTCCAGATANTAGAGACTGACCC 657

Qy   1289 GTCTTGTACAGTTTTTTTTTTTTTTTTTAATTGACTGTGGGAGTTTATGTTCCTCAGTTT 1348
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   658 AGTCTTGTACAGTTTNNTGTTTNT----AATGACTGGTGGGAGTTNATG-TCCTCTANGTT 711

Qy   1349 TTGTGAGGTTTGACTTAGTTATTATATAAATATTGTTTTTGTTGTTCATGTGA 1402
              ||||||||||||||||| | ||| || |||||| ||||
Db   712 TTGTGAGGTTTGACTAAGTTATAAATACCTGTNNTTGTTGNTNGNTCANGTGA 765

RESULT 14
AA816049
LOCUS   AA816049   639 bp   mRNA   linear   EST 13-FEB-1998
DEFINITION vr14h02.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
            IMAGE:1120659 5' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
            RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION   AA816049
VERSION   AA816049.1 GI:2885645
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 639)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE   The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
  COMMENT   Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.linl.gov) for further information.
            MGI:609995
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 510.
FEATURES   Location/Qualifiers
     source   1..639
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C3H"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1120659"
                     /cell_line="C2C12"
                     /lab_host="DH10B"
                     /clone_lib="Barstead mouse myotubes MPLRB5"
                     /note="Vector: pTT3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCGAAGTGGAGCGGCGGCCGCTTTTTTTTTTTTTTTTTTTTTT
                     3'], double-stranded cDNA was ligated to Eco RI adaptors
                     [AATTCGGATCCTTG], digested with Not I and cloned into the
                     Not I and Eco RI sites of the modified pTT3 vector.
                     Library constructed by Bob Barstead. The C2C12 cell line
                     (available from ATCC, catalog # CRL-1772) differentiates
                     rapidly, forming contractile myotubes and producing

RESULT 15
BX362477
LOCUS   BX362477   952 bp   mRNA   linear   EST 05-MAY-2003
DEFINITION BX362477 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ01SYP09 5-PRIME, mRNA sequence.
ACCESSION   BX362477
VERSION   BX362477.1 GI:30380640
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 952)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of

FEATURES   Location/Qualifiers
     source   1..952
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="CS0DJ015YP09"
                   /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                   /cell_line="JURKAT"
                   /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                   10-NORMALIZED"
                   /note="1st strand cDNA was primed with a Noti-oligo(dT)
                   primer. Five prime end enriched, double-strand cDNA was
                   digested with Not I and cloned into the Not I and EcoR V
                   sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match   33.1%; Score 621.2; DB 13; Length 952;
Best Local Similarity 82.5%; Pred. No. 4.7e-62;
Matches 738; Conservative 4; Mismatches 133; Indels 20; Gaps 2;

Qy   50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCGTGTAGAGC 109
          |||||||||| |||||| ||||| |||| |||||| |||| ||||| | |||||||
Db   68 GCAGGTAGCAAAGTGACGGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGAGAAC 127

Qy   110 GAGTGTTGCCATGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
          | | |||| |||||||||| |||||||||||| ||||||||||| ||| ||
Db   128 CAGCCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACGAGGA 182

Qy   170 AGTGGGGTCTGGAGACTATGACTCCAACAAGAACCCTGCTTCCGGGATGAAAAACGTCCA 229
          | ||||| ||| |||||||||||| ||| |||||| |||||| ||||||| | ||
Db   183 AATGGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA 242

Qy   230 TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCATCTTCTTGACTGGCATAGTCGG 289
          |||||||| | ||||||||||||||||| |||||||||| |||||||| |||||| || ||
Db   243 TTTCAATAAAATCTTCCTGCCCACCATCCATCCATCATCATCTTCTTAACTGGCATTGTGGG 302

Qy   290 CAATGGATTGGTGATCCTGGTCATGGTTACCACAAGAAGCTAAGGAGCATGACGGACAA 349
          |||||||||||| ||||||||||||| |||||| |||||| | ||| |||||||||||
Db   303 CAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGACGAAGCATGACGGACACA 362

Qy   350 GTACCGGCTGCACCTGTCAGTGGCTGACCCTCCTCTTGTCATCACACTCCCTTCTGGGC 409
          |||| |||||||||||||||||| || |||||||||||||||| || ||||||||
Db   363 GTACAGGCTGCACCTGTCAGTGGCGCCGACCTCCTCTTTGTCATCACGCTTCCCTTCTGGGC 422

Qy   410 AGTTGATGCCATGGCCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
          ||||||||||| ||| | ||||||| ||||||| || || ||||| |||||| ||||
Db   423 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGCAGTCCATGTCAT 482

Qy   470 CTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTA 529
          ||||| |||||||||||| |||| | |||||||||||||||||||| ||||||||| ||
Db   483 CTACACAGTCAACCTCTACGGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGCTA 542

Qy   530 CCTCGCCATTGTCCACGCCACCACCAAGTCAAAGGCAAGGAAAACTGCTGGCTGAAAAGGC 589
          ||| |||||||||||||||||||||||||||||||||||||| | ||||| |||||||| |
Db   543 CCTGGCCATTGTCCACGCCACCACCAAGTCAAAGGCAAGGAAGAAGCTGTTGGCTGAAAAGGT 602

Qy   590 AGTCTATGTGGGCGTCTGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGC 649
          ||||||||| |||||||| |||| |||| |||| ||||||| | |||||||||||||
Db   603 GGTCTATGTTGGCGTCTGGATCCCCTTCCCCTGCTGACTATTCCGACTTCATCTTTGC 662

Qy   650 CGACTGTCAGCCGACATCAGTGGACGACATCAGTGGGGATGCAGCAGTACACATCTGACCGCCTTA 709
          | | | | ||| ||||| | |||| | | | | |||||||| |||| ||||||| ||
Db   663 CAACG--------TCAGTANGGCAGATGACAGATATATCTGTGACCGCCTTCTA 707

Qy   710 CCCCGATAGCCTGTGGATGGTGGTTCAATTCCAGCAATAAATGGTGGGTCCTCATCCT 769
          |||| || ||| ||| |||| ||| | || |||| | |||||| ||||||||||
Db   708 CCCCAATGACTTGTGGGGTGGTTGTGTGTTCCAGTTTCAGCACATCATGGTTGGCCTTATCCT 767

Qy   770 GCCCGGCATCGTCATCCTCCTCCGTGTTACTGACGATCATCATCCTAAGCTGTCACACTCCAA 829
          ||| || ||||||||||| |||| || | ||| || | ||| ||| ||||||||||||||||
Db   768 GCCTGGGTATTGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAA 827

Qy   830 GGGCCACCAGAAGCGCAAGGCCCTCAAGACCAGTACAGCAGTCATCCTCATCCTAGCTTCTTTGC 889
          |||||||||||||||||||||||||||||||| | || ||||| || ||||||||| |||||| ||
Db   828 GGGCCACCAGAAGCGCAAGGCCCTCAAGACCACCAGTCATCCTCATCCTGGCTTTCTTCGC 887

Qy   890 CTGCTGGCTGCCATATTATGTGGGGATCAGCATCAGCATGACTCCTTCATCCTTTGGGA 944
          ||| |||||||| | || || ||||||||||| ||| ||| |||| ||||| | : ||
Db   888 CTGTTGGCTGCCTTACTACATTGGGATCAGCATGGGATCAGCATGCGACTCCCTTCATCYCYCTKGSAA 942

Search completed: May 17, 2004, 16:08:53
Job time : 5439.72 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2004, 16:09:21 ; Search time 134 Seconds
(without alignments)
756.974 Million cell updates/sec

Title: US-09-367-052-2
Perfect score: 1867
Sequence: 1 MEPISVSIYTSDNYSEEVGS.........KRGGHSSVSTBSESSSFHSS 359

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1867	100.0	359	2	AAW64778	Aaw64778 A murine
2	1867	100.0	359	2	AAY39994	Aay39994 Mouse CXC
3	1744.5	93.4	349	7	ADE63141	Ade63141 Rat Prote
4	1744.5	93.4	349	7	ADE63145	Ade63145 Rat Prote
5	1678.5	89.9	352	2	AAR68812	Aar68812 Human mon
6	1678.5	89.9	352	2	AAR80757	Aar80757 Human CXC
7	1678.5	89.9	352	2	AAY39993	Aay39993 Human CXC
8	1678.5	89.9	352	4	AAY52507	Aay52507 Human CXC
9	1678.5	89.9	352	4	AAE06690	Aae06690 Human neu
10	1678.5	89.9	352	4	AAG79088	Aag79088 Amino aci
11	1678.5	89.9	352	4	AAG80123	Aag80123 Human CXC
12	1678.5	89.9	352	5	AAO14003	Aao14003 Human cys
13	1678.5	89.9	352	5	ABP52651	Abp52651 Human CXC
14	1678.5	89.9	352	5	ABG32977	Abg32977 Human CXC
15	1678.5	89.9	352	5	ABG33065	Abg33065 Human CXC
16	1678.5	89.9	352	5	ABP65183	Abp65183 Hypoxia-r
17	1678.5	89.9	352	5	AAG78471	Aag78471 CXCR4 pro
18	1678.5	89.9	352	6	ABP96794	Abp96794 Human COP
19	1678.5	89.9	352	6	ABU03495	Abu03495 Angiogene
20	1678.5	89.9	352	6	ABR58581	Abr58581 Human can
21	1678.5	89.9	352	6	ABP97733	Abp97733 Amino aci
22	1678.5	89.9	352	6	ABP81796	Abp81796 Human CXC
23	1678.5	89.9	352	6	ADC98646	Adc98646 Human CXC
24	1678.5	89.9	352	7	ADE63147	Ade63147 Human Pro
25	1678.5	89.9	352	7	ADE63143	Ade63143 Human Pro
26	1672.5	89.6	352	4	ABB56346	Abb56346 Non-endog
27	1672.5	89.6	356	2	AAW97362	Aaw97362 G-protein
28	1670.5	89.5	352	6	ABU79148	Abu79148 Chemokine
29	1670.5	89.5	352	6	ABU79080	Abu79080 Human che
30	1667.5	89.3	352	2	AAR27792	Aar27792 New plate
31	1621.5	86.9	337	3	AAW48734	Aaw48734 Human RM3
32	1621.5	86.9	337	2	AAB21698	Aab21698 Human 7TM
33	1621.5	86.3	337	5	AAU91240	Aau91240 Human 7 t
34	1611.5	86.3	337	2	AAR53753	Aar53753 Seven tra
35	969.5	51.9	209	7	ADC86235	Adc86235 Human GPC
36	584	31.3	368	2	AAW54371	Aaw54371 Human IP-
37	584	31.3	368	4	AAY90614	Aay90614 Human G p
38	584	31.3	368	6	AAG80122	Aag80122 Human CXC
39	584	31.3	368	6	ABP81795	Abp81795 Human CXC
40	584	31.3	368	7	AAE38602	Aae38602 Human C-X
41	584	31.3	368	7	ADC22539	Adc22539 Human G p
42	584	31.3	378	2	AAY50129	Aay50129 Human che
43	584	31.3	415	7	ADC46963	Adc46963 Human CXC
44	581.5	31.1	367	2	AAW69999	Aaw69999 Rodent ch
45	580	31.1	368	3	AAY90648	Aay90648 Human mut

ALIGNMENTS

RESULT 1
AAW64778
ID   AAW64778 standard; protein; 359 AA.
XX
AC   AAW64778;
XX
DT   20-NOV-1998 (first entry)
XX
DE   A murine CXC chemokine receptor.
XX
KW   Mouse; CXC chemokine receptor; pre-B cell line DW34;
KW   CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1; HIV infection;
KW   screening; inhibitor; AIDS.
XX
OS   Mus sp.
XX
PN   WO983035-A1.
XX
PD   13-AUG-1998.
XX
PF   07-FEB-1997;   97WO-JP000299.
PR   07-FEB-1997;   97WO-JP000299.
XX
PA   (SHIO) SHIONOGI & CO LTD.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;
PI   Nakajima T, Yoshie O;
XX
DR   WPI; 1998-447232/38.
DR   N-PSDB; AAV46370.
XX
PT   Mouse CXC chemokine receptor binding to PBSF/SDF-1 pre-B cell stimulatory
PT   factor - is useful for screening of potential HIV infection and AIDS
PT   inhibitors.
XX
PS   Claim 1; Page 54-56; 76pp; Japanese.
XX
CC   The present sequence represents a murine CXC chemokine receptor which
CC   binds to the mouse CXC chemokine pre-B cell stimulatory factor PBSF/SDF-
CC   1. The nucleic acid is isolated from mouse pre-B cell line DW34. The
CC   receptor and cells expressing it can be used in the study and mapping of
CC   the mechanism of HIV infection and in screening of potential inhibitors
CC   of HIV infection and the development of AIDS
XX
SQ   Sequence 359 AA;

Query Match   100.0%; Score 1867; DB 2; Length 359;

CC   and tissue repairing agents containing as the active ingredient a
CC   substance capable of potentiating CXCR4. It is also possible to establish
CC   methods for treatment with the use of these remedies
XX
SQ   Sequence 359 AA;

     Query Match   100.0%; Score 1867; DB 2; Length 359;
     Best Local Similarity   100.0%; Pred. No. 9.7e-215;
     Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGL 60

Qy   61 VILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTV 120

Qy   121 NLYSSVLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVS 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NLYSSVLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVS 180

Qy   181 QGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQ 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 QGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQ 240

Qy   241 KRKALKTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFF 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 KRKALKTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFF 300

Qy   301 HCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 HCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359

RESULT 3
ADE63141
ID   ADE63141 standard; protein; 349 AA.
XX
AC   ADE63141;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Rat Protein O08565, SEQ ID NO 9076.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO) GEN HOSPITAL CORP.
PA   (FARB) BAYER AG.
XX
PI   Woolf C, D'urso D, Befort K, Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; O08565.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX

     Best Local Similarity   100.0%; Pred. No. 9.7e-215;
     Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGL 60

Qy   61 VILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTV 120

Qy   121 NLYSSVLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVS 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 NLYSSVLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVS 180

Qy   181 QGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQ 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 QGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQ 240

Qy   241 KRKALKTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFF 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 KRKALKTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFF 300

Qy   301 HCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 HCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359

RESULT 2
AAY39994
ID   AAY39994 standard; protein; 359 AA.
XX
AC   AAY39994;
XX
DT   16-DEC-1999 (first entry)
XX
DE   Mouse CXCR4 protein sequence.
XX
KW   CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;
KW   tissue repairing agent; vascularisation.
XX
OS   Mus sp.
XX
PN   WO9948528-A1.
XX
PD   30-SEP-1999.
XX
PF   23-MAR-1999; 99WO-JP001448.
XX
PR   24-MAR-1998; 98JP-00095448.
XX
PA   (CHUS) CHUGAI SEIYAKU KK.
PA   (KISH/) KISHIMOTO T.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K;
XX
DR   WPI; 1999-591042/50.
DR   N-PSDB; AAZ27611.
XX
PT   CXCR4-potentiating agents and methods useful for inhibiting
PT   neovascularization, and treating solid cancers.
XX
PS   Disclosure; Page 49-50; 63pp; Japanese.
XX
CC   This sequence is the mouse CXCR4 protein. The invention relates to
CC   remedies inhibiting neovascularisation, remedies for solid cancer,
CC   remedies for diseases pathologically caused by neovascularisation and
CC   tissue repairing agents containing as the active ingredient a substance
CC   capable of potentiating CXCR4. Based on a finding that vascularisation is
CC   inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC   remedies inhibiting vascularisation which contain as the active
CC   ingredient a substance capable of potentiating CXCR4, remedies for solid
CC   cancer, remedies for diseases pathologically caused by neovascularisation

Query Match 93.4%; Score 1744.5; DB 7; Length 349;
Best Local Similarity 95.5%; Pred. No. 4.7e-200;
Matches 338; Conservative 4; Mismatches 7; Indels 5; Gaps 1;

Qy   6 VSIYTSDNYSEEVGSGGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
         : |||||||||||||||||||||||||| :||||||||||||||||||||||||||||||
Db   1 MEIYTSDNYSEEVGSGGDYDSNKEPCFRDENENFNRIFLPTIYFIIFLTGIVGNGLVILVM 60

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 120

Qy 126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
         |||||||||||||||||||||||| ||||||||||||||||||||||||||----S 175
Db 121 VLILAFISLDRYLAIVHATNSQSARKLLAEKAVYVGVWIPALLLTIPDIIFADV----S 175

Qy 186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 176 QGDGRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 235

Qy 246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCFESIVHKWISITEALAFFHCCLN 305
         ||||||||||||||||||||||||||||||||||||:|:|||||||||||||||||||
Db 236 KTTVILILAFFACWLPYYVGISIDSFILLEVIKQGCEFESVVHKWISITEALAFFHCCLN 295

Qy 306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 349

RESULT 4
ADE63145
ID ADE63145 standard; protein; 349 AA.
AC ADE63145;
DT 29-JAN-2004 (first entry)
DE Rat Protein O08565, SEQ ID NO 9080.
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
OS Rattus norvegicus.

```
Qy   186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
          :|||||| ||: |:||:|||||||||||||||||||||||||||:||:|||||||||||||
Db   179  EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy   246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
          |||||||||||||||||||||||||||||:|||:||:|||||||||||||||||||||||
Db   239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy   306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352


RESULT 6
AAR80757
ID   AAR80757 standard; protein; 352 AA.
XX
AC   AAR80757;
XX
DT   25-MAR-2003  (revised)
DT   26-MAR-1996  (first entry)
XX
DE   Chemokine superfamily receptor.
XX
KW   Interleukin; IL-8; inflammation; psoriasis; dermatitis;
KW   rheumatoid arthritis; inflammatory bowel disease;
KW   chronic lung inflammation; treatment; antibody; affinity purification;
KW   detection.
XX
OS   Homo sapiens.
XX
PN   US5440021-A.
XX
PD   08-AUG-1995.
XX
PF   25-FEB-1994;     94US-00202056.
XX
PR   29-MAR-1991;     91US-00677211.
XX
PA   (HEBE/) HEBERT C.
PA   (CHUN/) CHUNTHARAPAI A.
PA   (KIMK/) KIM K J.
PA   (LEEJ/) LEE J.
XX
PI   Chuntharapai A,  Kim KJ,  Hebert C,   Lee J;
XX
DR   WPI; 1995-283151/37.
DR   N-PSDB; AAQ99007.
XX
PT   New antibodies against interleukin 8 type B receptor - used to treat or
PT   prevent inflammation. When immobilised, these antibodies may be used to detect
PT   interleukin-8 receptor B expression in cells and tissues and for affinity
PT   purification of interleukin-8 receptor B from cells. This sequence is an
PT   additonal chemokine superfamily receptor which was identified by probing
PT   lamda libraries of genomic DNA from a human monocyte-like cell line (L-
PT   60) and human peripheral blood lymphocytes using a large fragment of the
PT   interleukin-8 type A receptor DNA (See AAQ99006). (Updated on 25-MAR-2003
PT   to correct PF field.)
XX
PS   Example 2; Col 45-48; 62pp; English.
XX
CC   Antibodies directed against the interleukin-8 receptor B can be used to
CC   treat or prevent inflammation e.g. psoriasis, dermatitis, rheumatoid
CC   arthritis and particularly inflammatory bowel disease and chronic lung
CC   inflammation. When immobilised, these antibodies may be used to detect
CC   interleukin-8 receptor B expression in cells and tissues and for affinity
CC   purification of interleukin-8 receptor B from cells. This sequence is an
CC   additonal chemokine superfamily receptor which was identified by probing
CC   lamda libraries of genomic DNA from a human monocyte-like cell line (L-
CC   60) and human peripheral blood lymphocytes using a large fragment of the
CC   interleukin-8 type A receptor DNA (See AAQ99006). (Updated on 25-MAR-2003
CC   to correct PF field.)
XX
SQ   Sequence 352 AA;

Query Match          89.9%;  Score 1678.5;  DB 2;  Length 352;
Best Local Similarity 90.1%;  Pred. No. 4e-192;
Matches 319;   Conservative  19;  Mismatches  11;  Indels  5;  Gaps  1;
```

```
Qy   246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
          |||||||||||||||||||||||||||||||||||||||:|||:||||||||||||||||
Db   236  KTTVILILAFFACWLPYYVGISIDSFILLEVIKQGCEFESVVHKWISITEALAFFHCCLN  295

Qy   306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
          |||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db   296  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  349


RESULT 5
AAR68812
ID   AAR68812 standard; protein; 352 AA.
XX
AC   AAR68812;
XX
DT   25-MAR-2003  (revised)
DT   18-JUL-1995  (first entry)
XX
DE   Human monocyte PF4AR.
XX
KW   Interleukin-8 receptor; IL-8 receptor; PF4AR;
KW   platelet factor superfamily receptor; monocyte; chemotactic;
KW   inflammation; inflammatory disease; arthritis; emphysema; cystic;
KW   fibrosis; colitis; bronchitis; meningitis; therapeutic.
XX
OS   Homo sapiens.
XX
PN   WO9428931-A1.
XX
PD   22-DEC-1994.
XX
PF   07-JUN-1994;    94WO-US006380.
XX
PR   11-JUN-1993;    93US-00076093.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chuntharapai A,  Lee J,  Hebert C,  Kim KJ;
XX
DR   WPI; 1995-036114/05.
DR   N-PSDB; AAQ80521.
XX
PT   Treatment of inflammatory disorders - by administering an antibody
PT   capable of binding a platelet factor 4 superfamily receptor polypeptide.
XX
PS   Disclosure; Page 54-56; 83pp; English.
XX
CC   2 PF4AR members were identified by probing lambda libraries from human
CC   monocyte-like cell line HL-60 and human peripheral blood lymphocytes
CC   using a large fragment of IL-8 receptor DNA (full sequence given in
CC   AAQ80520). The nucleotide sequences of the 2 PF4ARs are given in AAQ80521
CC   and AAQ80522, and their respective amino acid sequences in AAR68812 and
CC   AAR68813. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 352 AA;

Query Match          89.9%;  Score 1678.5;  DB 2;  Length 352;
Best Local Similarity 90.1%;  Pred. No. 4e-192;
Matches 319;   Conservative  19;  Mismatches  11;  Indels  5;  Gaps  1;

Qy     6  VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM  65
          :|||||||||:|||||||||||:||||:|: :|||||||:|||||:|||||||||||||||
Db     4  ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM  63

Qy    66  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAWDWTFPGKFLCKAVHIIYTNLYSS  125
          ||||||||||||||||||||||||||||||||||||||:::|||||||||:||||||||
Db    64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy   126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYTVGVWIPALLLTIPDFIFADVSQGDIS  185
          ||||||:||||||||||||||||||||||||||:|||||||||||||||||:||:|
Db   124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVYVGVWIPALLLTIPDFIFANV-----S  178
```

XX
SQ   Sequence 352 AA;

Query Match   89.9%; Score 1678.5; DB 2; Length 352;
Best Local Similarity   90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
          :||||||||::|:|:||||||||||||||:||:::||:||||:||||||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          ||||||||||||||||||||||||||||||||||::|:||||||||||||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          ||||||||||||||||||||||||||||||:||||||||||||||||:||||:
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          ||||||||||:|:||||||||||||||||||||||||||||||||||||||||||||||
Db   179 EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISTEALAFFHCCLN 305
          ||||||||||||||||||||||||||||||:||||:||||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||:|||||:|||||||||||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

RESULT 8
AAY52507
ID   AAY52507 standard; protein; 352 AA.
XX
AC   AAY52507;
XX
DT   22-FEB-2000 (first entry)
XX
DE   Human CXC-chemokine receptor-4 (CXCR-4).
XX
KW   CXC-chemokine receptor-4; CXCR-4; stromal cell derived factor-1; SDF-1;
KW   ligand; chemotaxis; inflammation; G-protein-coupled receptor;
KW   signal transduction; CD4-independent; HIV-1; infection; proliferation;
KW   transformation; tumorigenesis; cancer; tumour; overexpression; brain;
KW   breast; colon; lung; melanoma; glioblastoma; inhibition; growth arrest;
KW   diagnosis; prognosis; marker; proliferative disorders; antisense;
KW   therapy; treatment; premalignant condition; hypertrophy;
KW   degenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   growth deficiency; hypoproliferative disorder; physical trauma; lesion;
KW   ischaemia; wound.
XX
OS   Homo sapiens.
XX
PN   WO9950461-A1.
XX
PD   07-OCT-1999.
XX
PF   29-MAR-1999;   99WO-US007431.
XX
PR   30-MAR-1998;   98US-0079916P.
PR   16-OCT-1998;   98US-0104656P.
XX
PA   (NWBI-) NORTHWEST BIOTHERAPEUTICS INC.
XX
PI   Murphy GP, Boynton AL, Sehgal A;
XX
DR   WPI; 2000-052634/04.
DR   N-PSDB; AAZ38553.
XX
PT   Use of CXCR-4 and SDF-1 as markers for diagnosis and treatment of e.g.
PT   tumors, degenerative disorders, growth deficiencies, hyper- and

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
          :||||||||::|:|:||||||||||||||:||:::||:|||||||||||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          ||||||||||||||||||||||||||||||||||::|:||||||||||||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          ||||||||||||||||||||||||||||||:||||||||||||||||:||||:
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          ||||||||||:|:||||||||||||||||||||||||||||||||||||||||||||||
Db   179 EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISTEALAFFHCCLN 305
          ||||||||||||||||||||||||||||||:||||:||||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||:|||||:|||||||||||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

RESULT 7
AAY39993
ID   AAY39993 standard; protein; 352 AA.
XX
AC   AAY39993;
XX
DT   16-DEC-1999 (first entry)
XX
DE   Human CXCR4 protein sequence.
XX
KW   CXCR4; human; neovascularisation; inhibitor; solid cancer; therapy;
KW   tissue repairing agent; vascularisation.
XX
OS   Homo sapiens.
XX
PN   WO9948528-A1.
XX
PD   30-SEP-1999.
XX
PF   23-MAR-1999;   99WO-JP001448.
XX
PR   24-MAR-1998;   98JP-00095448.
XX
PA   (CHUS) CHUGAI SEIYAKU KK.
PA   (KISH/) KISHIMOTO T.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K;
XX
DR   WPI; 1999-591042/50.
DR   N-PSDB; AAZ27610.
XX
PT   CXCR4-potentiating agents and methods useful for inhibiting
PT   neovascularization, and treating solid cancers.
XX
PS   Disclosure; Page 47-48; 63pp; Japanese.
XX
CC   This sequence is the human CXCR4 protein. The invention relates to
CC   remedies inhibiting neovascularisation, remedies for solid cancer,
CC   remedies for diseases pathologically caused by neovascularisation and
CC   tissue repairing agents containing as the active ingredient a substance
CC   capable of potentiating CXCR4. Based on a finding that vascularisation is
CC   inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC   remedies inhibiting vascularisation which contain as the active
CC   ingredient a substance capable of potentiating CXCR4, remedies for solid
CC   cancer, remedies for diseases pathologically caused by neovascularisation
CC   and tissue repairing agents containing as the active ingredient a
CC   substance capable of potentiating CXCR4. It is also possible to establish
CC   methods for treatment with the use of these remedies

PT   hypoproliferative disorders, physical trauma, lesions and wounds.
XX
PS   Disclosure; Fig 14; 138pp; English.
XX
CC   This sequence represents CXC-chemokine receptor-4 (CXCR-4). Chemokine
CC   receptors play an important role in the chemotaxis of T-cells and
CC   phagocytic cells to areas of inflammation. CXCR-4 is a member of the G-
CC   protein-coupled receptor family, which is involved in signal
CC   transduction, and its ligand is stromal cell derived factor-1 (SDF-1,
CC   AAY52508). CXCR-4 also mediates CD4-independent infection by HIV-1. CXCR-
CC   4 has now been found to have a role in the aberrant proliferative
CC   behaviour of a number of cell types, including numerous primary tumours
CC   and derived cell lines. CXCR-4 is involved in cell transformation and
CC   tumorigenesis, particularly in brain, breast and colon tumours. It was
CC   found to be overexpressed in several brain tumour derived cell lines and
CC   primary brain tumour tissues, breast tumour tissues, colorectal
CC   adenocarcinoma, lung carcinoma and melanoma cell lines. CXCR-4 expression
CC   was required for the continuous proliferation of glioblastoma cancer
CC   cells, and inhibition of its gene function resulted in growth arrest.
CC   Conversely, overexpression resulted in enhanced and rapid cellular
CC   proliferation. CXCR-4 and SDF-1 can be used as markers for the diagnosis
CC   and prognosis of cell proliferative disorders, and antisense
CC   oligonucleotides complementary to at least a portion of an RNA transcript
CC   of a CXCR-4 gene can be used to inhibit hyperproliferation of a tumour
CC   cell. Agents that inhibit CXCR-4 function can be used for treating or
CC   preventing a disease or disorder involving cell overproliferation, e.g.,
CC   brain cancer, breast cancer, colon cancer, prostate cancer and B cell
CC   lymphoma, and also premalignant conditions, benign tumours,
CC   hyperproliferative disorders, and benign dysproliferative disorders. They
CC   can also be used for treating e.g., cirrhosis of the liver, keloid
CC   formation, psoriasis, benign tumors, fibrocystic conditions and tissue
CC   hypertrophy. Compounds that promote CXCR-4 function can also be used for
CC   preventing or treating a disease or disorder involving a deficiency in
CC   cell proliferation, or treating a condition where cell proliferation
CC   would be desirable. Such diseases include degenerative disorders (e.g.,
CC   Parkinson's disease, Alzheimer's disease) growth deficiencies,
CC   hypoproliferative disorders, physical trauma, lesions (e.g., those caused
XX   by ischaemia), and wounds
SQ   Sequence 352 AA;

Query Match   89.9%; Score 1678.5; DB 3; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels   5; Gaps   1;

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
           :||||||||| ||||||||||:|||||||||:|::|:|||||||||:||||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
           ||||||||||||||||||||||||||||||||||:|| |||:||||||||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
           ||||||||||||||||||||||||||||||||:||||||||||||||||||||:
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
           :|||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db   179 EADDRYICDRFYPNDLMVVVFQFQHIMVGLILPGIVIKQGCDFESIVHKWISITBALAFFHCCLN 238

Qy   246 KTTVILIAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITBALAFFHCCLN 305
           :|||||||||||||||||||||||:||||:||||||||||||||||||||||||||||||
Db   239 KTTVILIAFFACWLPYYIGISIDSFILLEIKQGCFENTVHKWISITBALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
           ||||||||||||||:||||||||||||||||||||||||||||||||||||:::
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

RESULT 9
AAE06690
ID   AAE06690 standard; protein; 352 AA.

XX
AC   AAE06690;
XX
DT   16-OCT-2001 (first entry)
XX
DE   Human neuropeptide Y (NPY) Y3 receptor.
XX
KW   Human; neuropeptide Y; NPY; bone disease; bone mass; gene therapy;
KW   cerebrospinal fluid; CSF; inositol phosphate; IP; Paget's disease;
KW   fracture; extracellular signal-regulated kinase; ERK; osteoporosis;
KW   osteopenia; bone metastasis; neurotransmitter; osteogenic;
KW   NPY Y3 receptor.
XX
OS   Homo sapiens.
XX
PN   WO200153477-A1.
XX
PD   26-JUL-2001.
XX
PF   22-JAN-2001; 2001WO-US002040.
XX
PR   20-JAN-2000; 2000US-00489872.
XX
PA   (BAYU) BAYLOR COLLEGE MEDICINE.
PA   (AMLi/) AMLING M.
XX
PI   Amling M, Karsenty G, Ducy P;
XX
DR   WPI; 2001-488709/53.
DR   N-PSDB; AAD12801.
XX
PT   Treating or preventing reduced bone mass, e.g. osteoporosis, by reducing
PT   the level of neuropeptide Y activity in blood or cerebrospinal fluid.
XX
PS   Example 7; Page 85; 102pp; English.
XX
CC   The present invention relates to a method for treatment or prevention of
CC   bone diseases characterised by loss of bone mass, comprises administering
CC   to a mammal a compound that lowers the level of neuropeptide Y (NPY) in
CC   the serum or cerebrospinal fluid (CSF) or a compound that lowers the
CC   level of inositol phosphate (IP) or extracellular signal-regulated kinase
CC   (ERK). The method is specifically used to treat (including by gene
CC   therapy) or prevent osteoporosis, osteopenia or Paget's disease, but may
CC   also be used e.g., in cases of fractures or bone metastases. These
CC   diseases may also be diagnosed by detecting elevated NPY levels,
CC   including monitoring of treatment, assessing efficacy of compounds in
CC   clinical trials and for identifying subjects at risk. The present
CC   sequence is a human NPY Y3 receptor
XX
SQ   Sequence 352 AA;

Query Match   89.9%; Score 1678.5; DB 4; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels   5; Gaps   1;

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
           :||||||||| ||||||||||:|||||||||:|::|:||||||||:||||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
           ||||||||||||||||||||||||||||||||||:|| |||:||||||||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
           ||||||||||||||||||||||||||||||||:||||||||||||||||||||:
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
           :|||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db   179 EADDRYICDRFYPNDLMVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILIAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITBALAFFHCCLN 305
           :|||||||||||||||||||||||:||||:||||||||||||||||||||||||||||||
Db   246 KTTVILIAFFACWLPYYIGISIDSFILLEIKQGCFEREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM

Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          |||||||||||:||:|||:|||||:|||||:||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

XX

RESULT 10
AAG79088
ID   AAG79088 standard; protein; 352 AA.
XX
AC   AAG79088;
XX
DT   10-DEC-2001 (first entry)
XX
DE   Amino acid sequence of human CXCR4 protein.
XX
KW   Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;
KW   C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection;
KW   CXCR4.
XX
OS   Homo sapiens.
XX
PN   WO200164752-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US006322.
XX
PR   02-MAR-2000; 2000US-00517605.
XX
PA   (UYNY) UNIV NEW YORK STATE.
PA   (UYNI-) UNIV NIJMEGEN.
XX
PI   Littman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;
XX
DR   WPI; 2001-602565/68.
XX
PT   An antibody for the treatment or prevention of HIV-infection comprises a
PT   gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of
PT   DC-SIGN due to concomitant conformational change.
XX
PS   Disclosure; Page 117-118; 131pp; English.
XX
CC   The specification describes an antibody which is specific for an
CC   antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or
CC   is exposed upon gp120 binding of DC-SIGN due to concomitant
CC   conformational change. DC-SIGN is a receptor that is specifically
CC   expressed on dendritic cells and facilitates infection of T lymphocytes
CC   with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC
CC   -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)
CC   with high affinity. The antibody of the invention inhibits the trans
CC   enhancement of HIV entry into a T cell or macrophage facilitated by
CC   dendritic cells. The antibody is useful to treat or prevent HIV
CC   infection. The present sequence represents a human CXCR4 protein, which
CC   is a translocation promoting agent that interacts with CD4. This receptor
CC   functions in HIV-1 entry into cells
XX
SQ   Sequence 352 AA;

Query Match   89.9%; Score 1678.5; DB 4; Length 352;
Best Local Similarity   90.1%; Pred. No. 4e-192;
Matches   319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy   6 VSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLPLPTIYFIIPLLTGIVGNGLVILVM 65
          ||||||||||:||||||||||||||| ||||:||:||||||:|||:||||||||||||||||
Db   4 ISIYTSDNYTEMGSGGYDSMKEPCFREENANFNKIFLPLPTIYSIIFLFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          ||||||||||||||:||||||||||||||||||||:||||||:|||||||:|||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYYGVWIPALLLTIPDFIFADVSQGDIS 185

Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYYGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          : ||||||:::|:|::|||:||:||||||||||||||||||||||||:||||||||||||
Db   179 EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
          |||||||||||||||||||||||||||||||:||||:||||::|||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          |||||||||||||:|||||:||:|||:|||||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

XX

RESULT 11
AAG80123
ID   AAG80123 standard; protein; 352 AA.
XX
AC   AAG80123;
XX
DT   17-JAN-2002 (first entry)
XX
DE   Human CXCR4 protein.
XX
KW   Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
KW   inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
KW   chronic bowel inflammation; rheumatoid arthritis; cytostatic;
KW   antiinflammatory; antiasthmatic; immunosuppressive; dermatological;
KW   antirheumatic; antiarthritic.
XX
OS   Homo sapiens.
XX
PN   WO200172830-A2.
XX
PD   04-OCT-2001.
XX
PF   02-APR-2001; 2001WO-EP003708.
XX
PR   31-MAR-2000; 2000DE-01016013.
XX
PA   (IPFP-) IPF PHARM GMBH.
PA   (FORS/) FORSSMANN U.
XX
PI   Forssmann W, Adermann K, Heitland A, Spodsberg N;
XX
DR   WPI; 2001-626256/72.
XX
PT   Diagnostic agent containing two or more receptor-specific ligands, useful
PT   for detecting tumors, inflammation etc., also therapeutic use of ligand
PT   inhibitors.
XX
PS   Disclosure; Page 12; 26pp; German.
XX
CC   This invention describes a novel diagnostic agent (A) comprising at least
CC   two different ligands (I) for receptors (II) that are implicated in
CC   disease. (A) are used for the diagnosis of tumors (especially colorectal
CC   or prostatic), organ rejection, inflammation and autoimmune diseases.
CC   Also inhibitors of (I) are used therapeutically against tumors (and their
CC   metastases), inflammation (particularly bronchial asthma or chronic bowel
CC   inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
CC   where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
CC   endocrine, motor or urogenital systems or skin are affected, and bone
CC   marrow diseases. The products of the invention are chemokine derivatives
CC   which have cytostatic, antiinflammatory, antiasthmatic,
CC   immunosuppressive, dermatological, antirheumatic, antiarthritic.
CC   Chemokines act on specific tumor and inflammatory cells through a
CC   constellation of chemokine receptors (CR), which control migration and
CC   proliferation of these cells. AAG0045-AAG80128 represent human chemokine
CC   fragments used to illustrate the method of the invention
XX
SQ   Sequence 352 AA;

```
Query Match          89.9%;  Score 1678.5;  DB 4;  Length 352;
Best Local Similarity 90.1%;  Pred. No. 4e-192;
Matches  319;  Conservative  19;  Mismatches  11;  Indels  5;  Gaps  1;

Qy     6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVM 65
          ||||||:||||||||||||| |||||:|||||||||| ||:|||||||||||||||
Db     4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy    66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          |||||||||||||||||||||||||||||||||:|:|||||||:||||:||||||||
Db    64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          |||||||||||||||||||||||||||||:|||||||||||||||||||:|||:
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLMVVFQFQHIMVGLIILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          : ||||||||:|:||||||||||||||||||||||||||||||:|||||||||||||
Db   179 EADDRYICDRFYPNDLMVVFQFQHIMVGLIILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSPILLGVIKQGCDFPESIVHKWISITEALAFFHCCLN 305
          |||||||||||||||||:||||||||:|||| |:||||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSSFHSS 359
          |||||||||||:||||||:|||||||||||||||||||||||||:||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSSFHSS 352


RESULT 12
AAO14003
ID   AAO14003  standard; protein; 352 AA.
XX
AC   AAO14003;
XX
DT   08-MAY-2002  (first entry)
XX
DE   Human cysteine-X-cysteine chemokine receptor 4.
XX
KW   Cysteine-X-Cysteine Chemokine Receptor 4; CXCR4 binding peptide;
KW   chemoattractant cytokine; human immunodeficiency virus; HIV;
KW   HIV envelope glycoprotein; cellular CD4 receptor; env gene; gp160; gp120;
KW   gp41; CD4+ target cell; HIV-infected cell; CXCR4-binding compound;
KW   acquired immunodeficiency syndrome; AIDS; human; receptor.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FH
FT   Region          1..39
FT                   /note="Extracellular N-terminus; this is involved in HIV
FT                   entry into a cell"
FT   Region          40..64
FT                   /note="Transmembrane segment 1 (TM1)"
FT   Region          65..76
FT                   /note="Intracellular loop 1"
FT   Region          77..99
FT                   /note="Transmembrane segment 2 (TM2)"
FT   Region          100..110
FT                   /note="Extracellular loop 1"
FT   Region          111..131
FT                   /note="Transmembrane segment 3 (TM3)"
FT   Region          132..176
FT                   /note="Intracellular loop 2"
FT   Region          177..197
FT                   /note="Transmembrane segment 4 (TM4)"
FT   Region          198..203
FT                   /note="Extracellular loop 2; required for HIV entry into
FT                   a cell"
FT   Region          204..223
FT                   /note="Transmembrane segment 5 (TM5)"
FT   Region          224..240
FT                   /note="Intracellular loop 3"
```

```
FT   Region          241..261
FT                   /note="Transmembrane segment 6 (TM6)"
FT   Region          262..282
FT                   /note="Extracellular loop 3; this is involved in HIV
FT                   entry into a cell"
FT   Binding-site    275
FT                   /note= "Binds HIV gp120"
FT   Binding-site    278
FT                   /note= "Binds HIV gp120"
FT   Misc-difference 280..281
FT                   /note= "Apparently encoded by nucleotides 838-903 of
FT                   AAK9824l"
FT   Region          283..307
FT                   /note= "Transmembrane segment 7 (TM7)"
FT   Region          308..352
FT                   /note= "Intracellular C-terminus"
XX
PN   WO200170768-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009160.
XX
PR   21-MAR-2000; 2000US-0190946P.
PR   21-MAR-2000; 2000US-0190996P.
PR   21-MAR-2000; 2000US-0191299P.
PR   20-MAR-2001; 2001US-00813448.
PR   20-MAR-2001; 2001US-00813651.
PR   20-MAR-2001; 2001US-00813653.
XX
PA   (CONS-) CONSENSUS PHARM INC.
XX
PI   Nestor JJ, Wilson CJ, Tan Hehir CA, Kates SA;
XX
DR   WPI; 2002-010593/01.
DR   N-PSDB; AAK9824l.
XX
PT   Identifying agents that bind to CXC chemokine receptor 4, useful for
PT   treatment and prevention of acquired immunodeficiency syndrome, comprises
PT   screening compound library for members that bind to the receptor.
XX
PS   Example 3; Fig 3B; 46pp; English.
XX
CC   The invention relates to a method of identifying Cysteine-X-Cysteine
CC   chemokine receptor 4 (CXCR4) binding compounds. The method comprises
CC   testing and selecting compounds in a library for their ability to bind to
CC   CXCR4. Chemokines (chemoattractant cytokines) regulate the movement and
CC   biological activities of leukocytes in many disease situations. CXCR4
CC   acts as a co-receptor for human immunodeficiency virus (HIV), as it
CC   interacts with the HIV envelope glycoproteins and the cellular CD4
CC   receptor, to facilitate HIV entry into cells. The HIV glycoproteins
CC   (encoded by the env gene) are translated as a precursor (gp160) which is
CC   subsequently cleaved into gp120 and gp41. Gp120 binds to CXCR4 and the
CC   CD4 receptor present on the surface of susceptible CD4+ target cells,
CC   causing the virus to fuse with the cell membranes and thereby
CC   facilitating entry of the virus into the cell. The eventual expression of
CC   the env product on the surface of the HIV-infected cell enables the cell
CC   to fuse with uninfected CD4+ cells, thereby spreading the infection. The
CC   CXCR4-binding compounds of the invention inhibit the entry of HIV into
CC   the CD4+ cell and therefore are useful in the prevention of HIV infection
CC   and the treatment and prevention of acquired immunodeficiency syndrome
CC   (AIDS). The present sequence represents the human CXCR4 protein used in
CC   the method of the invention
XX
SQ   Sequence 352 AA;

Query Match          89.9%;  Score 1678.5;  DB 5;  Length 352;
Best Local Similarity 90.1%;  Pred. No. 4e-192;
Matches  319;  Conservative  19;  Mismatches  11;  Indels  5;  Gaps  1;

Qy     6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVM 65
          ||||||:|||||||||||||:|||||:|||||||||||||||:||||||||||||||||
Db     4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63
```

```
Qy   66  GYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
             ||||||||||||||||||||||||||||||:|:||||||||:||||||:||||||||||
Db   64  GYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVYGVWIPALLLTIPDFIFADVSQGDIS  185
             ||||||||||||||||||||||||||||||:||||||||||||||||||:|:
Db  124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYGVWIPALLLTIPDFIFANV-----S  178

Qy  186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
             :||||||||:||:|:|||||||||||||||||||||||||||||||||||||||||||
Db  179  EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
             |||||||||||||||||:|||||||||:::|||:|:|:|||||||||||||||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
             ||||||||||||:||||||:|:||||||||||||||||||||||||||||||||
Db  299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

RESULT 13
ABP52651
ID   ABP52651 standard; protein; 352 AA.
XX
AC   ABP52651;
XX
DT   30-OCT-2002   (first entry)
XX
DE   Human CXC chemokine receptor 4 protein SEQ ID NO:126.
XX
KW   Cysteine-X-cysteine chemokine receptor 4; CXC chemokine receptor 4;
KW   binding; identification; AIDS; HIV; CXCR4; virucide; anti-HIV.
XX
OS   Homo sapiens.
XX
PN   WO200257313-A2.
XX
PD   25-JUL-2002.
XX
PF   26-OCT-2001; 2001WO-US051165.
XX
PR   27-OCT-2000; 2000US-0243587P.
PR   20-MAR-2001; 2001US-00813448.
PR   20-MAR-2001; 2001US-00813651.
PR   20-MAR-2001; 2001US-00813653.
XX
PA   (CONS-) CONSENSUS PHARM INC.
XX
PI   Nestor JJ, Wilson CJ, Tan Hehir CA, Kates SA, Krstenansky J;
XX
DR   WPI: 2002-608378/65.
DR   N-PSDB: ABQ74938.
XX
PT   Identifying a binding compound for cysteine-X-cysteine chemokine receptor
PT   4 (CXCR4) for preventing or treating AIDS and HIV infection, comprises
PT   binding a molecule to a molecule with a binding property corresponding to
PT   CXCR4.
XX
PS   Example 3; Fig 3A-B; 92pp; English.
XX
CC   The present invention describes a method (M1) for identifying a binding
CC   compound for cysteine-X-cysteine chemokine receptor 4 (CXCR4) comprises
CC   binding a molecule from the library of two or more molecules to the
CC   molecule with a binding property corresponding to CXCR4. (M1) comprises:
CC   (a) providing a library of two or more molecules; (b) providing a
CC   molecule with a binding property corresponding to CXCR4; (c) binding a
CC   molecule from the library of two or more molecules to the molecule with a
CC   binding property corresponding to CXCR4; (d) separating the bound
CC   molecule from the library of two or more molecules; and (e) identifying
CC   the bound molecule as a binding compound for CXCR4. Also described is a
CC   binding compound (I) for CXCR4 identified from M1. (I) has virucide and
CC   anti-HIV activities. A partially purified CXCR4 protein is useful for
CC   carrying out selection, identification and improvement of tight binding
CC   ligands in identifying therapeutically useful compounds. Therapeutic
CC   peptides, peptidomimetics, or small molecule antagonists or agonists of
CC   CXCR4 binding may be used in the prevention and treatment of AIDS and HIV
CC   infection. The present sequence represents human CXCR4 which is given in
CC   an example from the present invention
XX
SQ   Sequence 352 AA;

Query Match   89.9%; Score 1678.5; DB 5; Length 352;
Best Local Similarity   90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

```
Qy    6  VSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM   65
             :||||||||:|:|||||||||:|||||:|::|:||||||||:||||||||||||||||
Db    4  ISIYTSDNYTEEWGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy   66  GYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
             ||||||||||||||||||||||||||||||:|:||||||||:||||||:||||||||||
Db   64  GYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVYGVWIPALLLTIPDFIFADVSQGDIS  185
             ||||||||||||||||||||||||||||||:|||||||||||||||||||:----:S
Db  124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYGVWIPALLLTIPDFIFANV-----S  178

Qy  186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVIKQGCDFESIVHKWISITEALAFFHCCLN  245
             :||||||||:||:|:|||||||||||||||||:::
Db  179  EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
             |||||||||||||||||:|||||||||:::|||:|:|:|||||||||||||||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
             ||||||||||||:||||||:|:||||||||||||||||||||||||||||||||
Db  299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

RESULT 14
ABG32977
ID   ABG32977 standard; protein; 352 AA.
XX
AC   ABG32977;
XX
DT   02-DEC-2002   (first entry)
XX
DE   Human CXC chemokine receptor 4 (CXCR4).
XX
KW   CXC chemokine receptor 4; CXCR4; chemotaxis induction; SDF1a;
KW   stromal cell-derived factor 1alpha; angiogenesis; atherosclerosis;
KW   restenosis; ischaemic stroke; spinal cord injury; infection; ulcer;
KW   human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;
KW   AIDS; pain; cancer; benign prostatic hypertrophy; diabetes; obesity;
KW   anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW   hypotension; hypertension; urinary retention; osteoporosis; stroke;
KW   anginal pectoris; myocardial infarction; benign prostatic hypertrophy;
KW   migraine; vomiting; psychotic disorder; neurological disorder; anxiety;
KW   schizophrenia; dyskinesia; Huntingdon's disease; restenosis;
KW   inflammatory disease; rheumatoid arthritis; diabetic retinopathy;
KW   inflammatory bowel disease; atherosclerosis; Alzheimer's disease;
KW   congestive heart failure; cardiac remodeling; angiogenic diseases;
KW   solid tumour; Kaposi sarcoma; human.
XX
OS   Homo sapiens.
XX
PN   US2002107195-A1.
XX
PD   08-AUG-2002.
XX
PF   17-SEP-2001; 2001US-00955692.
XX
PR   21-JUL-1998; 98US-0093596P.

Left entry

```
PR   21-JUL-1999;   99US-00358524.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Gupta SK;
XX
DR   WPI; 2002-697879/75.
DR   N-PSDB; ABS53992.
XX
PT   Inducing chemotaxis of endothelial cells, useful for regulating
PT   angiogenesis, e.g. for treating cancer, comprises treatment with stromal
PT   cell-derived factor 1 alpha.
XX
PS   Claim 3; Fig 1; 26pp; English.
XX
CC   The invention describes a method of inducing chemotaxis of endothelial
CC   cells by treatment with stromal cell-derived factor 1alpha (SDFia). The
CC   method is used for stimulating EC chemotaxis, and thus angiogenesis, and
CC   is used for treating atherosclerosis, restenosis, ischaemic stroke and
CC   spinal cord injury. Inhibition of this process is useful in treatment and
CC   prevention of a very wide range of diseases, such as, infection (by
CC   bacteria, fungi, protozoa or viruses such as human immunodeficiency virus
CC   (HIV) and acquired immunodeficiency syndrome (AIDS)), pain, cancer and
CC   benign prostatic hypertrophy, diabetes, obesity, anorexia, bulimia,
CC   asthma, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, anginal pectoris,
CC   myocardial infarction, stroke, ulcers, benign prostatic hypertrophy,
CC   migraine, vomiting, psychotic and neurological disorders (e.g. anxiety,
CC   schizophrenia) and dyskinesias (e.g. Huntingdon's disease), inflammatory
CC   diseases, rheumatoid arthritis, diabetic retinopathy, inflammatory bowel
CC   disease, atherosclerosis, restenosis, Alzheimer's disease, congestive
CC   heart failure, cardiac remodeling, angiogenic diseases, solid tumours,
CC   and Kaposi Sarcoma. This is the amino acid sequence of a CXC chemokine
CC   receptor 4 (CXCR4) polynucleotide
XX
SQ   Sequence 352 AA;

Query Match             89.9%;    Score 1678.5;   DB 5;   Length 352;
Best Local Similarity   90.1%;    Pred. No. 4e-192;
Matches 319;   Conservative   19;   Mismatches 11;   Indels   5;   Gaps   1;

Qy     6  VSIYTSDNYSBEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLPTGIVGNGLVILVM   65
             :|||||||| :||||||||:|||||:|: :|:|||||||||::|||||||||||||||
Db     4  ISIYTSDNYTBEMGSGDYDSMKEPCFRRENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy    66  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
          |||||||||||||||||||||||||||||||||:|: ||||:||||||||:||||||||
Db    64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy   126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
          |||||||||||||||||||||||||||||| :|||||||||||||||||||::
Db   124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S  178

Qy   186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
          :.|||||||| :|| .||:|||||||||||||||||||||||||||||||||||||||||
Db   179  EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy   246  KTVILLAFFACWLPYYVGISDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
          ||||||||||||||||:|||:||||| :|||||:||:.|||||||||||||||||||
Db   239  KTVILLAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy   306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESBSSSFHSS  359
          ||||||||||||:||||||.|| ||||||||||||||||||||||||||||||
Db   299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESBSSSFHSS  352

RESULT 15
ABG33065
ID   ABG33065 standard; protein; 352 AA.
XX
AC   ABG33065;
XX
DT   28-NOV-2002  (first entry)
```

Right entry

```
XX
DE   Human CXCR4 protein.
XX
KW   Chemotaxis; endothelial cell; EC; angiogenesis; atherosclerosis;
KW   restenosis; ischaemic stroke; spinal cord injury; infection; bacteria;
KW   fungi; protozoa; virus; pain; cancer; benign prostatic hypertrophy;
KW   diabetes; obesity; anorexia; bulimia; asthma; allergy; hypertension;
KW   Parkinson's disease; acute heart failure; hypotension; urinary retention;
KW   osteoporosis; angina pectoris; myocardial infarction; stroke; dyskinesia;
KW   migraine; vomiting; psychotic disorder; neurological disorder; ulcer;
KW   inflammatory disorder; rheumatoid arthritis; diabetic retinopathy;
KW   inflammatory bowel disease; atherosclerosis; restenosis; Kaposi sarcoma;
KW   Alzheimer's disease; congestive heart failure; cardiac remodelling;
KW   angiogenic disease; solid tumour; human; stromal cell derived factor-la;
KW   SDF-1alpha; CXCR4.
XX
OS   Homo sapiens.
XX
PN   US2002107196-A1.
XX
PD   08-AUG-2002.
XX
PF   17-SEP-2001; 2001US-00953717.
XX
PR   21-JUL-1998;   98US-0093596P.
PR   21-JUL-1999;   99US-00358624.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Gupta SK;
XX
DR   WPI; 2002-706230/76.
DR   N-PSDB; ABS53750.
XX
PT   Inducing chemotaxis of endothelial cells, useful for regulating
PT   angiogenesis, e.g. for treating cancer, comprises treatment with stromal
PT   cell-derived factor 1 alpha.
XX
PS   Claim 3; Fig 1; 26pp; English.
XX
CC   The present invention relates to a new method for inducing chemotaxis of
CC   endothelial cells (EC). The method of the invention involves treatment
CC   with a stromal cell-derived factor 1a (SDFialpha). The method is used for
CC   stimulating EC chemotaxis, and thus angiogenesis, and is used for
CC   treating atherosclerosis, restenosis, ischaemic stroke and spinal cord
CC   injury, while inhibition of this process is useful in treatment and
CC   prevention of a very wide range of diseases, e.g. infection (by bacteria,
CC   fungi, protozoa or viruses), pain, cancer and benign prostatic
CC   hypertrophy, diabetes, obesity, anorexia, bulimia, asthma, allergies,
CC   Parkinson's disease, acute heart failure, hypotension, hypertension,
CC   urinary retention, osteoporosis, angina pectoris, myocardial infarction,
CC   stroke, ulcers, migraine, vomiting, psychotic and neurological disorders
CC   and dyskinesias, inflammatory disorders, rheumatoid arthritis, diabetic
CC   retinopathy, inflammatory bowel disease, atherosclerosis, restenosis,
CC   Alzheimer's disease, congestive heart failure, cardiac remodelling,
CC   angiogenic diseases, solid tumours, and Kaposi sarcoma. The present amino
CC   acid sequence represents the human CXCR4 protein of the invention
XX
SQ   Sequence 352 AA;

Query Match             89.9%;    Score 1678.5;   DB 5;   Length 352;
Best Local Similarity   90.1%;    Pred. No. 4e-192;   English.
Matches 319;   Conservative   19;   Mismatches 11;   Indels   5;   Gaps   1;

Qy     6  VSIYTSDNYSBEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLPTGIVGNGLVILVM   65
             :|||||||| :||||||||:|||||:|: :|:|||||||||::|||||||||||||||
Db     4  ISIYTSDNYTBEMGSGDYDSMKEPCFRREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy    66  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
          |||||||||||||||||||||||||||||||||:|: ||||:||||||||:||||||||
Db    64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy   126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
```

```
Db  124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYGVWIPALLLTIPDFIFANV-----S  178
              |||||||||||||||||||||||||||||||||||||||       ||||||||||||||||||:|

Qy  186  QGDDRYICDRLYPDSLWMVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
              :|||||||||||: ||:||||||||||||||||||||||||||||||||||||||||
Db  179  EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
              ||||||||||||||||||:||||||||||: |||::|||||| |||||||||||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIIKQGCBFENTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
              |||||||||||:||||||| |::|||||||||||||||||||||||||||
Db  299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

Search completed: May 17, 2004, 22:16:14
Job time : 141 secs

OM protein - protein search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: May 17, 2004, 22:14:28 ; Search time 31 Seconds
(without alignments)
597.862 Million cell updates/sec

Title: US-09-367-052-2
Perfect score: 1867
Sequence: 1 MEPISVSIYTSDNYSEEVGS.........KRGGHSSVSTESESSSFHSS 359

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1678.5	89.9	352	1	US-08-202-056-3	Sequence 3, Appli
2	1678.5	89.9	352	1	US-08-076-093A-4	Sequence 4, Appli
3	1678.5	89.9	352	1	US-08-701-265-4	Sequence 4, Appli
4	1678.5	89.9	352	2	US-08-284-586-4	Sequence 4, Appli
5	1678.5	89.9	352	2	US-08-805-478-4	Sequence 4, Appli
6	1678.5	89.9	352	2	US-08-802-627A-4	Sequence 4, Appli
7	1678.5	89.9	352	2	US-08-801-238-4	Sequence 4, Appli
8	1678.5	89.9	352	2	US-08-801-228-4	Sequence 4, Appli
9	1678.5	89.9	352	4	US-09-104-296-4	Sequence 4, Appli
10	1678.5	89.9	352	4	US-09-517-605-4	Sequence 4, Appli
11	1641.5	87.9	352	1	US-08-450-393A-6	Sequence 6, Appli
12	1641.5	87.9	352	3	US-08-446-669-6	Sequence 6, Appli
13	1641.5	87.9	352	3	PCT-US95-00476-6	Sequence 6, Appli
14	1621.5	86.9	337	1	US-08-153-848-46	Sequence 46, Appl
15	1621.5	86.9	337	4	US-09-299-843A-46	Sequence 46, Appl
16	1621.5	86.9	337	4	US-09-088-337B-46	Sequence 46, Appl
17	1621.5	86.9	337	5	PCT-US93-11153-46	Sequence 46, Appl
18	584	31.3	368	3	US-08-709-838-2	Sequence 2, Appli
19	584	31.3	368	3	US-08-829-839-2	Sequence 2, Appli
20	584	31.3	368	4	US-09-170-496D-20	Sequence 20, Appli
21	584	31.3	368	4	US-09-624-594-2	Sequence 2, Appli
22	580	31.1	368	4	US-09-170-496D-174	Sequence 174, App
23	570.5	30.6	360	3	US-08-875-573-20	Sequence 20, Appli
24	570.5	30.6	360	2	US-09-232-878-2	Sequence 2, Appli
25	570.5	30.6	360	3	US-09-048-583-55	Sequence 55, Appl
26	570.5	30.6	360	4	US-09-534-185-55	Sequence 55, Appl
27	570.5	30.6	360	4	US-08-939-107-34	Sequence 34, Appl
28	565.5	30.3	123	3	US-08-513-974B-330	Sequence 330, App
29	558.5	29.9	360	4	US-08-833-752-10	Sequence 10, Appl
30	550.5	29.5	378	4	US-09-299-843A-66	Sequence 66, Appl
31	550.5	29.5	378	4	US-09-088-337B-66	Sequence 66, Appli
32	544.5	29.2	355	1	US-07-759-568-1	Sequence 1, Appli
33	544.5	29.2	355	2	US-08-450-393A-8	Sequence 8, Appli
34	544.5	29.2	355	2	US-08-390-000A-5	Sequence 5, Appli
35	544.5	29.2	355	3	US-08-446-669-8	Sequence 8, Appli
36	544.5	29.2	355	5	PCT-US95-00476-8	Sequence 8, Appli
37	544.5	29.2	360	1	US-08-202-056-7	Sequence 7, Appli
38	544.5	29.2	360	4	US-09-409-778-4	Sequence 4, Appli
39	537.5	28.8	358	3	US-08-153-848-19	Sequence 19, Appli
40	537.5	28.8	358	3	US-09-299-843A-19	Sequence 19, Appl
41	537.5	28.8	358	4	US-09-088-337B-19	Sequence 19, Appl
42	537.5	28.8	358	5	PCT-US93-11153-19	Sequence 19, Appl
43	537.5	28.8	378	1	US-08-153-848-15	Sequence 15, Appl
44	537.5	28.8	378	3	US-09-299-843A-15	Sequence 15, Appl
45	537.5	28.8	378	3	US-09-251-545-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-202-056-3
; Sequence 3, Application US/08202056
; Patent No. 5440021
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Hebert, Caroline
; APPLICANT: Kim, Kyung Jin
; APPLICANT: Lee, James
; TITLE OF INVENTION: Antibodies to Human IL-8 Type B Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,056
; FILING DATE: 25-FEB-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-202-056-3

Query Match 89.9%; Score 1678.5; DB 1; Length 352;
Best Local Similarity 90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy 6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNLVILVM 65

```
Best Local Similarity 30.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy    6  VSIYTSDNYSEBVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM   65
            :|||||||| :||||||||:|||||||:: ::|||||||:||||||||||||||||||
Db    4  ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy   66  GYQKLRSWTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
            ||||||||||||||||||||||||||||||||||:::|||:|||||||:|:||||||
Db   64  GYQKLRSWTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIVTVNLYSS  123

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
            |||||||||||||||||||||||||||||:|||||||||||||||||||||:
Db  124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S  178

Qy  186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
            :||||||||:||:|:|:|||||||||||||||||||||||||||||||||||||||||
Db  179  EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
            ||||||||||||||||||:||||||||||  :|||||:|:  ||||||||||||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFNTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
            |||||||||||:||||||| :|:||||||||||||||||||||||||||||||
Db  299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

RESULT 3
US-08-701-265-4
; Sequence 4, Application US/08701265
; Patent No. 5776457
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,265
; FILING DATE: 22-AUG-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-Jun-1993
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

```
            :|||||||| :||||||||:|||||||:: ::|||||||:||||||||||||||||||
Db    4  ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy   66  GYQKLRSWTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
            ||||||||||||||||||||||||||||||||||:::|||:|||||||:|:||||||
Db   64  GYQKLRSWTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIVTVNLYSS  123

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
            |||||||||||||||||||||||||||||:|||||||||||||||||||||:
Db  124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S  178

Qy  186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
            :||||||||:||:|:|:|||||||||||||||||||||||||||||||||||||||||
Db  179  EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
            ||||||||||||||||||:||||||||||  :|||||:|:  ||||||||||||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFNTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
            |||||||||||:||||||| :|:||||||||||||||||||||||||||||||
Db  299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

RESULT 2
US-08-076-093A-4
; Sequence 4, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,093A
; FILING DATE: 11-Jun-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
US-08-076-093A-4
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear

Query Match 89.9%; Score 1678.5; DB 1; Length 352;

;   LENGTH: 352 amino acids
;   TYPE: Amino Acid
;   TOPOLOGY: Linear
US-08-701-265-4

Query Match   89.9%; Score 1678.5; DB 1; Length 352;
Best Local Similarity   90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
              :||||||| |||||||||||| :|||::||||| |||||||||| |||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFRRENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGSKFLCKAVHIIYTVNLYSS 125
              ||||||||||||||||||||||||||||||||: | :|||||||||||:: |||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
              |||||||||||||||||||||||||||||||: |: |: |||||||||||||:| |::
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
              |||||||||||||| |: |: |:||||||||||||||||||||||||: |||||||||||
Db   179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
              |||||||||||||||||||||||||||||:|::||:|: :|||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
              ||||||||||||:|||||| :|:|||||||||||||||||||||||:|||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHS 352

RESULT 4
US-08-284-586-4
; Sequence 4, Application US/08284586
; Patent No. 5840856
; GENERAL INFORMATION:
;   APPLICANT: Chuntharapai, Anan
;   APPLICANT: Lee, James
;   APPLICANT: Hebert, Caroline
;   APPLICANT: Jin Kim, K.
;   TITLE OF INVENTION: Antibodies to Human PF4A Receptors
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/284,586
;   FILING DATE:
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/076,093A
;   FILING DATE: 11-Jun-1993
;   APPLICATION NUMBER: 07/810782
;   FILING DATE: 19-DEC-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/677211
;   FILING DATE: 29-MAR-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: 706P2

;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 352 amino acids
;   TYPE: Amino Acid
;   TOPOLOGY: Linear
US-08-284-586-4

Query Match   89.9%; Score 1678.5; DB 2; Length 352;
Best Local Similarity   90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
              :||||||| |||||||||||| :|||::||||| |||||||||| |||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFRRENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
              ||||||||||||||||||||||||||||||||: | :|||||||||||:: |||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
              |||||||||||||||||||||||||||||||: |: |: |||||||||||||:| |::
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
              |||||||||||| |: |: |:||||||||||||||||||||||||: |||||||||||
Db   179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
              |||||||||||||||||||||||||||||:|::||:|: :|||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
              ||||||||||||:|||||| :|:|||||||||||||||||||||||:|||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHS 352

RESULT 5
US-08-805-478-4
; Sequence 4, Application US/08805478
; Patent No. 5874543
; GENERAL INFORMATION:
;   APPLICANT: Chuntharapai, Anan
;   APPLICANT: Lee, James
;   APPLICANT: Hebert, Caroline
;   APPLICANT: Jin Kim, K.
;   TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/805,478
;   FILING DATE: 25-Feb-1997
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/284586
;   FILING DATE: 10-AUG-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/076093
;   FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/810782
  FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Love, Richard B.
  REGISTRATION NUMBER: 34,659
  REFERENCE/DOCKET NUMBER: P0706P2P1C1
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415/225-5530
  TELEFAX: 415/952-9881
  TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
  LENGTH: 352 amino acids
  TYPE: Amino Acid
  TOPOLOGY: Linear
US-08-805-478-4

Query Match   89.9%;   Score 1678.5; DB 2;   Length 352;
Best Local Similarity   90.1%;   Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLPTIGVNGLVILVM 65
          :||||||:||::||||| |||||||:|::|| :||:|:||||||||:|||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFRRENANFNKIFLPTIYSIIFLPTIGVNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          |||||||||||||||||||||||||||||| ::|:||||||||:|||||||:||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy 126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          |||||||||||||||||||||||||||||||:| :|||||||||||||||
Db 124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy 186 QGDDRYICDRLYPDSLMMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          :|||||:|::||:|| ||||||||||||||||||||||||||||||||||||||||||||
Db 179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy 246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
          ||||||||||||||||||:| |||||||||||:| :|||||||||||||||||||||||
Db 239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy 306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||:|::|||| :||||||||:||||||||||||||||||||||
Db 299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

RESULT 6
US-08-802-627A-4
; Sequence 4, Application US/08802627A
; Patent No. 5892017
GENERAL INFORMATION:
  APPLICANT: Lee, James
  APPLICANT: Wood, William I.
  TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
  NUMBER OF SEQUENCES: 6
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Genentech, Inc.
    STREET: 460 Point San Bruno Blvd
    CITY: South San Francisco
    STATE: California
    COUNTRY: USA
    ZIP: 94080
  COMPUTER READABLE FORM:
    MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: WinPatin (Genentech)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/802,627A
    FILING DATE: 19-Feb-1997
    CLASSIFICATION: 514
  PRIOR APPLICATION DATA:

        APPLICATION NUMBER: 08/284586
        FILING DATE: 10-AUG-1994
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 08/076093
        FILING DATE: 11-JUN-1993
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 07/810782
        FILING DATE: 19-DEC-1991
    ATTORNEY/AGENT INFORMATION:
      NAME: Love, Richard B.
      REGISTRATION NUMBER: 34,659
      REFERENCE/DOCKET NUMBER: P0706P2P1D2
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: 415/225-5530
      TELEFAX: 415/952-9881
      TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
  LENGTH: 352 amino acids
  TYPE: Amino Acid
  TOPOLOGY: Linear
US-08-802-627A-4

Query Match   89.9%;   Score 1678.5; DB 2;   Length 352;
Best Local Similarity   90.1%;   Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLPTIGVNGLVILVM 65
          :||||||:||::||||| |||||||:|::|| :||:|:||||||||:|||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFRRENANFNKIFLPTIYSIIFLPTIGVNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          |||||||||||||||||||||||||||||| ::|:||||||||:|||||||:||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy 126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          |||||||||||||||||||||||||||||||:| :|||||||||||||||
Db 124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy 186 QGDDRYICDRLYPDSLMMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          :|||||:|::||:|| ||||||||||||||||||||||||||||||||||||||||||||
Db 179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy 246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
          ||||||||||||||||||:| |||||||||||:| :|||||||||||||||||||||||
Db 239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy 306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||:|::|||| :||||||||:||||||||||||||||||||||
Db 299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

RESULT 7
US-08-801-238-4
; Sequence 4, Application US/08801238
; Patent No. 5919896
GENERAL INFORMATION:
  APPLICANT: Lee, James
  APPLICANT: Wood, William I.
  TITLE OF INVENTION: PF4A RECEPTOR
  NUMBER OF SEQUENCES: 6
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Genentech, Inc.
    STREET: 460 Point San Bruno Blvd
    CITY: South San Francisco
    STATE: California
    COUNTRY: USA
    ZIP: 94080
  COMPUTER READABLE FORM:
    MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: WinPatin (Genentech)

    ZIP: 94080
COMPUTER READABLE FORM:
    MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/801,228
    FILING DATE: 19-Feb-1997
    CLASSIFICATION: 435
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/284586
    FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/076093
    FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 07/810782
    FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
    NAME: Love, Richard B.
    REGISTRATION NUMBER: 34,659
    REFERENCE/DOCKET NUMBER: P0706P2P1D3
TELECOMMUNICATION INFORMATION:
    TELEPHONE: 415/225-5530
    TELEFAX: 415/952-9881
    TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 352 amino acids
    TYPE: Amino Acid
    TOPOLOGY: Linear
US-08-801-228-4

Query Match   89.9%;   Score 1678.5;   DB 2;   Length 352;
Best Local Similarity   90.1%;   Pred. No. 6.6e-133;
Matches 319;   Conservative 19;   Mismatches 11;   Indels 5;   Gaps 1;

```
Qy    6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
         :||||||||:|:|||||| |||||| :||:||:|||||||||||||||||||||||||||
Db    4 ISIYTSDNYTEEMSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
         |||||||||||||||||||||||||||||||||::|::||||||||:||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
         ||||||||||||||||||||||||||||:||:|::||||||||||||||||| :
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
         |||||||||:|||: :|||||||||||||||||||||||||||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
         ||||||||||||||||||:||||||||||:||||:|||||||||||||||||||||||||
Db  239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
         ||||||||||:|:||||:|||:||||||||||||:||||||||||||||||||
Db  299 PILYAFLGAKFKISAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 9
US-09-104-296-4
; Sequence 4, Application US/09104296
; Patent No. 6087475
GENERAL INFORMATION:
    APPLICANT: Lee, James
    APPLICANT: Wood, William I.
    TITLE OF INVENTION: PF4A Receptors
    NUMBER OF SEQUENCES: 6
    CORRESPONDENCE ADDRESS:

CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/801,238
    FILING DATE: 19-Feb-1997
    CLASSIFICATION: 530
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/284586
    FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/076093
    FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 07/810782
    FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
    NAME: Love, Richard B.
    REGISTRATION NUMBER: 34,659
    REFERENCE/DOCKET NUMBER: P0706P2P1D1
TELECOMMUNICATION INFORMATION:
    TELEPHONE: 415/225-5530
    TELEFAX: 415/952-9881
    TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 352 amino acids
    TYPE: Amino Acid
    TOPOLOGY: Linear
US-08-801-238-4

Query Match   89.9%;   Score 1678.5;   DB 2;   Length 352;
Best Local Similarity   90.1%;   Pred. No. 6.6e-133;
Matches 319;   Conservative 19;   Mismatches 11;   Indels 5;   Gaps 1;

```
Qy    6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
         :||||||||:|:|||||| |||||| :||:||:|||||||||||||||||||||||||||
Db    4 ISIYTSDNYTEEMSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
         |||||||||||||||||||||||||||||||||::|::||||||||:||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
         ||||||||||||||||||||||||||||:||:|::||||||||||||||||| :
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
         |||||||||:|||: :|||||||||||||||||||||||||||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
         ||||||||||||||||||:||||||||||:||||:|||||||||||||||||||||||||
Db  239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
         ||||||||||:|:||||:|||:||||||||||||:||||||||||||||||||
Db  299 PILYAFLGAKFKISAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 8
US-08-801-228-4
; Sequence 4, Application US/08801228
; Patent No. 5922541
GENERAL INFORMATION:
    APPLICANT: Lee, James
    APPLICANT: Wood, William I.
    TITLE OF INVENTION: METHODS FOR DETECTION AND AMPLIFICATION OF
    TITLE OF INVENTION: PF4A RECEPTOR NUCLEIC ACID
    NUMBER OF SEQUENCES: 6
    CORRESPONDENCE ADDRESS:
    ADDRESSEE: Genentech, Inc.
    STREET: 460 Point San Bruno Blvd
    CITY: South San Francisco
    STATE: California
    COUNTRY: USA

```
; Sequence 4, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
;   APPLICANT: Littman, Dan R.
;   APPLICANT: Kwon, Douglas S.
;   APPLICANT: van Kooyk, Yvette
;   APPLICANT: Geijtenbeck, Tneo
;   TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
;   TITLE OF INVENTION: CELLS
;   FILE REFERENCE: 1049-1-017
;   CURRENT APPLICATION NUMBER: US/09/517,605
;   CURRENT FILING DATE: 2000-03-02
;   NUMBER OF SEQ ID NOS: 17
;   SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 4
;   LENGTH: 352
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-517-605-4

  Query Match           89.9%;  Score 1678.5;  DB 4;  Length 352;
  Best Local Similarity 90.1%;  Pred. No. 6.6e-133;
  Matches 319;  Conservative  19;  Mismatches  11;  Indels   5;  Gaps    1;

Qy      6  VSIYTSDNYSEEVGSGDYDSNKEPCPRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM   65
           :||||||||||:|||||||::||| ||||:::||| :|:|||||||||:|||||||||||
Db      4  ISIYTSDNYTEEMGSGDYDSMKEPCPREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy     66  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
           ||||||||||||||||||||||||||||||||||||::|||||||||||||||||||||
Db     64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy    126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db    124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S  178

Qy    186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
           :||||||||||||:| :|||||||||||||||||||||||||||||||||||||||||||
Db    179  EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy    246  KTTVILILAFFACWLPYYVGISIDSPILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
           |||||||||||||||||||||||||||| :|||||||:||||||||||||||||||||||
Db    239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy    306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
           ||||||||||||:||||||||||::|||||||||||||||||||||||||||||
Db    299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352


RESULT 11
US-08-450-393A-6
; Sequence 6, Application US/08450393A
; Patent No. 5707815
; GENERAL INFORMATION:
;   APPLICANT: Charo, Israel
;   APPLICANT: Coughlin, Shaun
;   TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
;   TITLE OF INVENTION: PROTEIN RECEPTORS
;   NUMBER OF SEQUENCES:  14
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;     STREET:  5 Palo Alto Square
;     CITY:  Palo Alto
;     STATE:  California
;     COUNTRY:  USA
;     ZIP:  94306-2155
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0,  Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/450,393A
```

```
;   ADDRESSEE: Genentech, Inc.
;   STREET:  1 DNA Way
;   CITY:  South San Francisco
;   STATE:  California
;   COUNTRY: USA
;   ZIP:  94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/104,296
;     FILING DATE:  24-June-1998
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/701265
;     FILING DATE:  22-AUG-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/664228
;     FILING DATE:  06-JUN-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/076093
;     FILING DATE:  11-JUN-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 07/810782
;     FILING DATE:  19-DEC-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Love, Richard B.
;     REGISTRATION NUMBER:  34,659
;     REFERENCE/DOCKET NUMBER: P0706P2C2
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  415/225-5530
;     TELEFAX:  415/952-9881
;     TELEX:  910/371-7168
; INFORMATION FOR SEQ ID NO:  4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 352 amino acids
;     TYPE:  Amino Acid
;     TOPOLOGY:  Linear
US-09-104-296-4

  Query Match           89.9%;  Score 1678.5;  DB 3;  Length 352;
  Best Local Similarity 90.1%;  Pred. No. 6.6e-133;
  Matches 319;  Conservative  19;  Mismatches  11;  Indels   5;  Gaps    1;

Qy      6  VSIYTSDNYSEEVGSGDYDSNKEPCPRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM   65
           :||||||||||:|||||||::||| ||||:::||| :|:|||||||||:|||||||||||
Db      4  ISIYTSDNYTEEMGSGDYDSMKEPCPREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy     66  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
           ||||||||||||||||||||||||||||||||||||::|||||||||||||||||||||
Db     64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy    126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db    124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S  178

Qy    186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
           :||||||||||||:| :|||||||||||||||||||||||||||||||||||||||||||
Db    179  EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy    246  KTTVILILAFFACWLPYYVGISIDSPILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
           |||||||||||||||||||||||||||| :|||||||:||||||||||||||||||||||
Db    239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy    306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
           ||||||||||||:||||||||||::|||||||||||||||||||||||||||||
Db    299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352


RESULT 10
US-09-517-605-4
```

```
        FILING DATE: May 25, 1995
        CLASSIFICATION: 424
        ATTORNEY/AGENT INFORMATION:
        NAME: Cserr, Luann
        REGISTRATION NUMBER: 31,822
        REFERENCE/DOCKET NUMBER: UCAL-237/02US
        TELECOMMUNICATION INFORMATION:
        TELEPHONE: 415-843-5165
        TELEFAX: 415-8857-0663
        TELEX: 380816CooleyPA
    INFORMATION FOR SEQ ID NO: 6:
        SEQUENCE CHARACTERISTICS:
        LENGTH: 352 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein
        HYPOTHETICAL: NO
US-08-450-393A-6

Query Match        87.9%;  Score 1641.5;  DB 1;  Length 352;
Best Local Similarity 88.7%;  Pred. No. 8.2e-130;
Matches 314;  Conservative 18;  Mismatches 17;  Indels 5;  Gaps 1;

Qy    6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM  65
          :|||||||| ||:|||||||:||||||| :||  ||:|||| ||.||| |||||||||||
Db    4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPYIYSIIFLYGIVGNGLVILVM  63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          ||||||||||||||||||||||||||||||||||| :|.|:| ||||||||:|||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          |||||||||||||||||||||||||||||||.||||||||||||||||||||.----- :
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
           .|||||||| ||..||.||||||||||||||||||||| ||||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILFCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
          |  |||||||||||||||:||||||||||: :|||||.|: ||||||||||||||||||
Db  239 KYYVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||.|||:||:|||||||||||||||||||||||||||||||||
Db  299 PILYAFLGAKFKYSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 12
US-08-446-669-6
; Sequence 6, Application US/08446669
; Patent No. 6132987
GENERAL INFORMATION:
   APPLICANT: Charo, Israel
   APPLICANT: Coughlin, Shaun
   TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
   TITLE OF INVENTION: PROTEIN RECEPTORS
   NUMBER OF SEQUENCES: 14
   CORRESPONDENCE ADDRESS:
     ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
     STREET: 5 Palo Alto Square
     CITY: Palo Alto
     STATE: California
     COUNTRY: USA
     ZIP: 94306-2155
   COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
     SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/446,669
     FILING DATE: May 25, 1995

```
        CLASSIFICATION: 435
        ATTORNEY/AGENT INFORMATION:
        NAME: Neeley, Richard
        REGISTRATION NUMBER: 30,092
        REFERENCE/DOCKET NUMBER: UCAL-237/01US
        TELECOMMUNICATION INFORMATION:
        TELEPHONE: 415-843-5000
        TELEFAX: 415-857-0663
        TELEX: 380816CooleyPA
    INFORMATION FOR SEQ ID NO: 6:
        SEQUENCE CHARACTERISTICS:
        LENGTH: 352 amino acids
        TYPE: amino acid
        TOPOLOGY: linear
        MOLECULE TYPE: protein
        HYPOTHETICAL: NO
US-08-446-669-6

Query Match        87.9%;  Score 1641.5;  DB 3;  Length 352;
Best Local Similarity 88.7%;  Pred. No. 8.2e-130;
Matches 314;  Conservative 18;  Mismatches 17;  Indels 5;  Gaps 1;

Qy    6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM  65
          :|||||||| ||:|||||||:||||||| :||  ||:|||| ||.||| |||||||||||
Db    4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPYIYSIIFLYGIVGNGLVILVM  63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          ||||||||||||||||||||||||||||||||||| :|.|:| ||||||||:|||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          |||||||||||||||||||||||||||||||.||||||||||||||||||||.----- :
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
           .|||||||| ||..||.||||||||||||||||||||| ||||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILFCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
          |  |||||||||||||||:||||||||||: :|||||.|: ||||||||||||||||||
Db  239 KYYVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||.|||:||:|||||||||||||||||||||||||||||||||
Db  299 PILYAFLGAKFKYSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 13
PCT-US95-00476-6
; Sequence 6, Application PC/TUS9500476
GENERAL INFORMATION:
   APPLICANT: The Regents of the University of California
   TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
   TITLE OF INVENTION: PROTEIN RECEPTORS
   NUMBER OF SEQUENCES: 14
   CORRESPONDENCE ADDRESS:
     ADDRESSEE: Robbins, Berliner & Carson
     STREET: 201 N. Figueroa Street, 5th Floor
     CITY: Los Angeles
     STATE: California
     COUNTRY: USA
     ZIP: 90012-2628
   COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
     SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER: PCT/US95/00476
     FILING DATE:
     CLASSIFICATION:
   ATTORNEY/AGENT INFORMATION:
     NAME: Berliner, Robert

; REGISTRATION NUMBER: 20,121
; REFERENCE/DOCKET NUMBER: 5555-291
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-977-1001
; TELEFAX: 310-977-1003
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
PCT-US95-00476-6

Query Match   87.9%; Score 1641.5; DB 5; Length 352;
Best Local Similarity 88.7%; Pred. No. 8.2e-130;
Matches 314; Conservative 18; Mismatches 17; Indels 5; Gaps 1;

```
Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
         :||||||||| :|| ||||||| :||||: :|: |||||||| :||| ||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLYGIVGNGLVILVM 63

Qy  66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
        ||||||||||||||||||||||||||||||||||||:|| :|||:||||||||:|||||||
Db  64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy 126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYYGVWIPALLLTIPDFIFADVSQGDIS 185
        ||||||||||||||||||||||||||||||||:||||||||||||||||||||:||----:
Db 124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYYGVWIPALLLTIPDFIFANV-----S 178

Qy 186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          |||||||||:||:| |:|||||||||||||||||||:|||||||||||||||||||||
Db 179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILFCYCIIISKLSHSKGHQKRKAL 238

Qy 246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
        |  |||||||||||||||:|||||||||| :|||||:||:|:||||||||||||||||||
Db 239 KYYVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy 306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
        ||||||||||||:|||| | ||||||||||||||||||||||||||||||||||
Db 299 PILYAFLGAKFKYSAQHLTVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 14
US-08-153-848-46
; Sequence 46, Application US/08153848
; Patent No. 5759804
; GENERAL INFORMATION:
;   APPLICANT: Godiska, Ronald
;   APPLICANT: Gray, Patrick W.
;   APPLICANT: Schweikart, Vicki L.
;   TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
;   NUMBER OF SEQUENCES: 64
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;   ADDRESSEE: Bicknell
;   STREET: 6300 Sears Tower, 233 South Wacker Drive
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60606
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/153,848
;   FILING DATE:
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/977,452

;   FILING DATE: 17-NOV-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: No. 5759804and, Greta E.
;   REGISTRATION NUMBER: 35,302
;   REFERENCE/DOCKET NUMBER: 31794
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (312) 474-6300
;   TELEFAX: (312) 474-0448
;   TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 46:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 337 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-153-848-46

Query Match   86.9%; Score 1621.5; DB 1; Length 337;
Best Local Similarity 90.4%; Pred. No. 3.7e-128;
Matches 309; Conservative 17; Mismatches 11; Indels 5; Gaps 1;

```
Qy  18 VGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDK 77
        :||||||| :|| ||||||: :|: |||||||| :||||||||||||||||||||||||||
Db   1 MGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVMGYQKKLRSMTDK 60

Qy  78 YRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRY 137
        ||||||||||||||||||||||||:|| :|||:||||||||:||||||||||||||||||||
Db  61 YRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVLILAFISLDRY 120

Qy 138 LAIVHATNSQRPRKLLAEKAVYYGVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLY 197
        ||||||||||||||||||||:||||||||||||||||||:||----     |||||||:||
Db 121 LAIVHATNSQRPRKLLAEKVVYYGVWIPALLLTIPDFIFANV-----SEADDRYICDRFY 175

Qy 198 PDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFA 257
        |:|:| |:||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db 176 PNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFA 235

Qy 258 CWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFK 317
        ||||||:|||||||||| :|||||:||:|:|||||||||||||||||||||||||||||||
Db 236 CWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLNPILYAFLGAKFK 295

Qy 318 SSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
        :|||| | ||||||||||||||||||||||||||||||||||
Db 296 TSAQHLTVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 337
```

RESULT 15
US-09-299-843A-46
; Sequence 46, Application US/09299843A
; Patent No. 6107475
; GENERAL INFORMATION:
;   APPLICANT: Godiska, Ronald
;   APPLICANT: Gray, Patrick W.
;   APPLICANT: Schweikart, Vicki L.
;   TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
;   NUMBER OF SEQUENCES: 66
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;   ADDRESSEE: Borun
;   STREET: 6300 Sears Tower, 233 South Wacker Drive
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60606
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/299,843A
;   FILING DATE:
;   CLASSIFICATION: 435

```
Query Match        86.9%;   Score 1621.5;   DB 3;   Length 337;
Best Local Similarity   90.4%;   Pred. No. 3.7e-128;
Matches 309;   Conservative 17;   Mismatches 11;   Indels 5;   Gaps 1;

Qy   18  VGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDK  77
         : |||||| ||||:||  : :|::|||::|||  :||||||||||||||:||||||||||
Db    1  MGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVMGYQKKLRSMTDK  60

Qy   78  YRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHIIYTVNLYSSVLILAFISLDRY  137
         |||||||||||||||||||||||||:|:||| :|||||||||||||||||||||||||||
Db   61  YRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVLILAFISLDRY  120

Qy  138  LAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLY  197
         ||||||||||||||||||||||||||||||||||||||:| :|:|||    ||||||
Db  121  LAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV----SEADDRYICDRPY  175

Qy  198  PDSLMWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFA  257
         |:| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  176  PNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFA  235

Qy  258  CWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFK  317
         |||||:|||||||||||||:|||||:||: :|||||||||||||||||||||||||||||
Db  236  CWLPYYIGISIDSFILLEIKQGCEFENTVHKWISITEALAFFHCCLNPILYAFLGAKFK  295

Qy  318  SSAQHALNSMSRGGSSLKILSKGKRGGHSSVSTESBESSSFHSS  359
         :|||||| |:|||||:||||||||||||||||||||||||||||
Db  296  TSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  337
```

Search completed: May 17, 2004, 22:22:03
Job time : 37 secs

OM protein - protein search, using sw model

```
                  GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:  May 17, 2004, 22:19:12 ; Search time 547 Seconds
                                   (without alignments)
                         182.625 Million cell updates/sec

Title:         US-09-367-052-2
Perfect score: 1867
Sequence:      1 MEPISVSIYTSDNYSEEVGS..........KRGGHSSVSTESESSSFHSS 359

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      1145568 segs, 278261457 residues

Total number of hits satisfying chosen parameters:    1145568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Published_Applications_AA:*
  1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
  2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
  3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
  4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
  5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
  6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
  7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
  8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
  9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1678.5	89.9	352	9	US-09-953-692-2	Sequence 2, Appli
2	1678.5	89.9	352	9	US-09-953-717-2	Sequence 2, Appli
3	1678.5	89.9	352	9	US-09-104-063-4	Sequence 4, Appli
4	1678.5	89.9	352	12	US-10-211-462-81	Sequence 81, Appl
5	1678.5	89.9	352	12	US-10-666-689-4	Sequence 4, Appl
6	1678.5	89.9	352	12	US-10-181-906-10	Sequence 10, Appl
7	1678.5	89.9	352	12	US-09-813-651B-85	Sequence 85, Appl
8	1678.5	89.9	352	12	US-10-151-274-4	Sequence 4, Appli
9	1678.5	89.9	352	12	US-10-170-385-331	Sequence 331, App
10	1678.5	89.9	352	14	US-10-225-567A-76	Sequence 76, Appl
11	1678.5	89.9	352	14	US-10-245-850-1	Sequence 1, Appli
12	1678.5	89.9	352	14	US-10-251-703-38	Sequence 38, Appl
13	1678.5	89.9	352	14	US-10-021-660-123	Sequence 123, App
14	1678.5	89.9	352	14	US-10-014-322A-126	Sequence 126, App
15	1678.5	89.9	352	14	US-10-239-423-79	Sequence 79, Appl
16	1678.5	89.9	352	15	US-10-160-401-3	Sequence 3, Appli
17	1678.5	89.9	352	15	US-10-341-434-178	Sequence 178, App
18	1678.5	89.9	352	15	US-10-372-683-2	Sequence 2, Appl
19	1678.5	89.9	352	15	US-10-440-464-78	Sequence 78, Appl
20	1678.5	89.6	352	15	US-10-452-015-1	Sequence 1, Appli
21	1672.5	89.6	352	11	US-09-826-509-485	Sequence 485, App
22	1670.5	89.5	352	9	US-09-870-759-37	Sequence 37, Appl
23	1670.5	89.5	352	9	US-09-870-759-144	Sequence 144, App
24	1670.5	89.5	352	10	US-09-751-708A-37	Sequence 37, Appl
25	1670.5	89.5	352	10	US-09-751-708A-144	Sequence 144, App
26	969.5	51.9	209	15	US-10-292-798-688	Sequence 688, App
27	964.5	51.7	209	14	US-10-017-161-790	Sequence 790, App
28	584	31.3	368	12	US-10-411-284-4	Sequence 4, Appli
29	584	31.3	368	14	US-10-251-385-20	Sequence 20, Appl
30	584	31.3	368	14	US-10-225-567A-74	Sequence 74, Appl
31	584	31.3	368	14	US-10-345-680-59	Sequence 59, Appl
32	584	31.3	368	14	US-10-251-686-2	Sequence 2, Appli
33	584	31.3	368	14	US-10-239-423-78	Sequence 78, Appl
34	584	31.3	368	15	US-10-295-027-752	Sequence 752, App
35	584	31.1	415	12	US-10-411-284-2	Sequence 2, Appli
36	580	31.1	368	14	US-10-251-385-174	Sequence 174, App
37	575	30.8	472	14	US-10-106-698-6402	Sequence 6402, Ap
38	570.5	30.6	360	9	US-09-837-446-2	Sequence 2, Appl
39	570.5	30.6	360	9	US-09-796-744-17	Sequence 17, Appl
40	570.5	30.6	360	9	US-09-764-413-20	Sequence 20, Appl
41	570.5	30.6	360	12	US-10-039-659-16	Sequence 16, Appl
42	570.5	30.6	360	13	US-10-120-394-20	Sequence 20, Appl
43	570.5	30.6	360	14	US-10-225-567A-66	Sequence 66, Appl
44	570.5	30.6	360	14	US-10-164-649-55	Sequence 55, Appl
45	570.5	30.6	360	14	US-10-231-452-48	Sequence 48, Appl

ALIGNMENTS

```
RESULT 1
US-09-953-692-2
; Sequence 2, Application US/09953692
; Patent No. US20020107195A1
; GENERAL INFORMATION:
; APPLICANT: Shalley, Gupta K.
; TITLE OF INVENTION: Method for Inducing Chemotaxis in Endothelial Cells by
; TITLE OF INVENTION: Administering Stromal Cell Derived Factor-1(
; FILE REFERENCE: P50676C1
; CURRENT APPLICATION NUMBER: US/09/953,692
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 09/358,624
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 60/093,596
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Human
US-09-953-692-2

Query Match             89.9%;  Score 1678.5;  DB 9;  Length 352;
Best Local Similarity   90.1%;  Pred. No. 1.1e-145;
Matches  319;  Conservative   19;  Mismatches   11;  Indels    5;  Gaps    1;

Qy          6 VSIYTSDNYSEEVGSGDYDSNKPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
              :||:||::||::|::||||||||:|:||||::|||||||||||||||||||||||||||
Db          4 ISIYTSDNYTEEMGSGDYDSMKEPCFRENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy         66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHIIYTVNLYSS 125
              ||||||||||||||||||||||||||||||||||:|::|||||||||:|||||||||||
Db         64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy        126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
              ||||||||||||||||||||||||||||||:|:||||||||||||||||||:|:   ||
Db        124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178
```

TITLE OF INVENTION: PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/104,063
FILING DATE: 24-June-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/701265
FILING DATE: 22-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/664228
FILING DATE: 06-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-104-063-4

Query Match   89.9%;   Score 1678.5; DB 9; Length 352;
Best Local Similarity   90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

```
Qy   6 VSIYTSDNYSEEVGSGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM 65
         :|||||||| ||:|||| ||||:|||||||:::.:.|||||||||:||||||||||||||||
Db   4 ISIYTSDNYTEEMGSGGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy  66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
        |||||||||||||||||||||||||||||||||||:.:.||:|||:||||||:||||||||
Db  64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy 126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
        ||||||||||||||||||||||||||||||:.::|||||||||||||||||||||:----- 
Db 124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVVGVWIPALLLTIPDFIFANV-----S 178

Qy 186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
        :.||||||||:||:.||.||||||||||||||||||||||||||||||||||||||||||
Db 179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy 246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
        ||||||||||||||||||:|||||||||| :||||:||.:|||||||||||||||||||||
Db 239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy 306 PILYAFLGAKFKSSAQHALNSMSRGGSLKILSKGKRGGHSSVSTESESSSFHSS 359
        |||||||||||| |:|||||| |:||.||||||||||||||||||||||||||||
Db 299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 2
US-09-953-717-2
; Sequence 2, Application US/09953717
; Patent No. US2002010719A1
; GENERAL INFORMATION:
; APPLICANT: Shalley, Gupta K.
; TITLE OF INVENTION: Method for Inducing Chemotaxis in Endothelial Cells by
; TITLE OF INVENTION: Administering Stromal Cell Derived Factor-1(
; FILE REFERENCE: P5067601
; CURRENT APPLICATION NUMBER: US/09/953,717
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 09/358,624
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 60/093,596
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Human
US-09-953-717-2

Query Match   89.9%;   Score 1678.5; DB 9; Length 352;
Best Local Similarity   90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

```
Qy   6 VSIYTSDNYSEEVGSGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM 65
         :|||||||| ||:|||| ||||:|||||||:::.:.|||||||||:||||||||||||||||
Db   4 ISIYTSDNYTEEMGSGGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy  66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
        |||||||||||||||||||||||||||||||||||:.:.||:|||:||||||:||||||||
Db  64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy 126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
        ||||||||||||||||||||||||||||||:.::|||||||||||||||||||||:-----
Db 124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVVGVWIPALLLTIPDFIFANV-----S 178

Qy 186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
        :.||||||||:||:.||.||||||||||||||||||||||||||||||||||||||||||
Db 179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy 246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
        ||||||||||||||||||:|||||||||| :||||:||.:|||||||||||||||||||||
Db 239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy 306 PILYAFLGAKFKSSAQHALNSMSRGGSLKILSKGKRGGHSSVSTESESSSFHSS 359
        |||||||||||| |:|||||| |:||.||||||||||||||||||||||||||||
Db 299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 3
US-09-104-063-4
; Sequence 4, Application US/09104063
; Patent No. US2002016356A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.

RESULT 4
US-10-211-462-81
; Sequence 81, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-81

```
Query Match             89.9%;  Score 1678.5;  DB 12;  Length 352;
Best Local Similarity   90.1%;  Pred. No. 1.1e-145;
Matches  319;  Conservative   19;  Mismatches   19;  Indels    5;  Gaps     1;

Qy    6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM  65
         :||||||:|||:|||:||:||||:|||: :|||||||||:||||||||||||||||||
Db    4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM  63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
        ||||||||||||||||||||||||||||| :||||:|::|||:||||||:||:|||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVVGVWIPALLLTIPDFIFADVSQGDIS 185
        |||||||||||||||||||||||||||||:||||||||||||||:|||||||:|
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVVGVWIPALLLTIPDFIFANV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
        :||||||:|||:|||||:|||||||||||||||||:||||||||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
        ||||||||||||||||||:|||||||:||| :||:||||:||:||||||||||||||||||
Db  239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
        |||||||||||| :|||:|:||||||||||||||||||||||||||||||||||
Db  299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 5
US-10-666-689-4
; Sequence 4, Application US/10666689
; Publication No. US20040037830A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Human PF4A Receptors, Nucleic Acid Encoding and
; TITLE OF INVENTION: Antibodies Binding Thereto
; FILE REFERENCE: P0706P2C2D2C1
; CURRENT APPLICATION NUMBER: US/10/666,689
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 09/104,063
; PRIOR FILING DATE: 1988-06-24
; PRIOR APPLICATION NUMBER: US 08/701,265
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/664,228
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/076,093
; PRIOR FILING DATE: 1993-06-11
; PRIOR APPLICATION NUMBER: US 07/810,782
; PRIOR FILING DATE: 1991-12-19
; PRIOR APPLICATION NUMBER: US 07/677,211
; PRIOR FILING DATE: 1991-03-29
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-666-689-4

```
Query Match             89.9%;  Score 1678.5;  DB 12;  Length 352;
Best Local Similarity   90.1%;  Pred. No. 1.1e-145;
Matches  319;  Conservative   19;  Mismatches   11;  Indels    5;  Gaps     1;

Qy    6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM  65
         :||||||:|||:|||:||:||||:|||: :|||||||||:||||||||||||||||||
Db    4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM  63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
        ||||||||||||||||||||||||||||| :||||:|::|||:||||||:||:|||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVVGVWIPALLLTIPDFIFADVSQGDIS 185
        |||||||||||||||||||||||||||||:||||||||||||||:|||||||:|
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVVGVWIPALLLTIPDFIFANV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
        :||||||:|||:|||||:|||||||||||||||||:||||||||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
        ||||||||||||||||||:|||||||:||| :||:||||:||:||||||||||||||||||
Db  239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
        |||||||||||| :|||:|:||||||||||||||||||||||||||||||||||
Db  299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 6
US-10-181-906-10
; Sequence 10, Application US/10181906
; Publication No. US20040053864A1
; GENERAL INFORMATION:
; APPLICANT: Karsenty, Gerard
; APPLICANT: Amling, Michael
; APPLICANT: Ducy, Patricia
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BONE FORMATION VIA
; TITLE OF INVENTION: MODULATION OF NEUROPEPTIDE Y ACTIVITY
; FILE REFERENCE: 9142-020-999
; CURRENT APPLICATION NUMBER: PCT/US01/02040
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US/10/181,906
; PRIOR FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-181-906-10

```
Query Match             89.9%;  Score 1678.5;  DB 12;  Length 352;
Best Local Similarity   90.1%;  Pred. No. 1.1e-145;
```

```
Matches 319;  Conservative 19;  Mismatches 11;  Indels 5;  Gaps 1;

Qy    6  VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM   65
         :|||||||||:|:||||||||||||||| :|: .:|||:||:|||||||||||||||||
Db    4  ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy   66  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
         |||||||||||||||||||||||||||||||||||:.||:||:||:|||||||||||
Db   64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
         ||||||||||||||||||||||||||||||  ||||||||||||||||||||||:|
Db  124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV----S  178

Qy  186  QGDDRYICDRLYPDSLMWVVFQFQHIMVGLLLPGIVILSCYCIIISKLSHSKGHQKRKAL  245
         :.||||||:||:||||| :||||||||||||||:||||||||||||||||||||||||
Db  179  EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
         |||||||||||||||||||||||||:|||||||||||  :||||:||||:|||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALNSMRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
         ||||||||||||||:||:||||||||:|||||||||||||||||||||||||
Db  299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

```
Query Match     89.9%;  Score 1678.5;  DB 12;  Length 352;
Best Local Similarity  90.1%;  Pred. No. 1.1e-145;
Matches 319;  Conservative 19;  Mismatches 11;  Indels 5;  Gaps 1;
```

```
Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
         |||||||||||||||||||||||||:|||||||||||  :||||:||||:|||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALNSMRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
         ||||||||||||||:||:||||||||:|||||||||||||||||||||||||
Db  299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

```
Query Match     89.9%;  Score 1678.5;  DB 12;  Length 352;
Best Local Similarity  90.1%;  Pred. No. 1.1e-145;
Matches 319;  Conservative 19;  Mismatches 11;  Indels 5;  Gaps 1;

Qy    6  VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM   65
         :|||||||||:|:||||||||||||||| :|: .:|||:||:|||||||||||||||||
Db    4  ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy   66  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVADWYFGKFLCKAVHIIYTVNLYSS  125
         ||||||||||||||||||||||||||||||||||||:.||:||:||:|||||||||||
Db   64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
         ||||||||||||||||||||||||||||||  ||||||||||||||||||||||:|
Db  124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV----S  178

Qy  186  QGDDRYICDRLYPDSLMWVVFQFQHIMVGLLLPGIVILSCYCIIISKLSHSKGHQKRKAL  245
         :.||||||:||:||||| :||||||||||||||:||||||||||||||||||||||||
Db  179  EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
         |||||||||||||||||||||||||:|||||||||||  :||||:||||:|||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALNSMRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
         ||||||||||||||:||:||||||||:|||||||||||||||||||||||||
Db  299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

```
;   APPLICANT: White, Jonathan
;   APPLICANT: Binley, Katie Mary
;   APPLICANT: Rayner, William Nigel
;   APPLICANT: Naylor, Stuart
;   APPLICANT: Kingsman, Susan Mary
;   APPLICANT: Krige, David
;   TITLE OF INVENTION: ANALYSIS METHOD
;   FILE REFERENCE: 53268200010 0
;   CURRENT APPLICATION NUMBER: US/10/170,385
;   CURRENT FILING DATE: 2002-06-12
;   PRIOR APPLICATION NUMBER: PCT/GB02/01662
;   PRIOR FILING DATE: 2002-04-08
;   PRIOR APPLICATION NUMBER: PCT/GB01/05458
;   PRIOR FILING DATE: 2001-12-10
;   NUMBER OF SEQ ID NOS: 549
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   SEQ ID NO 331
;   LENGTH: 352
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-170-385-331

  Query Match             89.9%;  Score 1678.5;  DB 12; Length 352;
  Best Local Similarity   90.1%;  Pred. No. 1.1e-145;
  Matches  319;  Conservative   19;  Mismatches   11;  Indels    5;  Gaps     1;

Qy          6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM 65
              :|||||||||:||:||||||||:||||||:||.|.|||||||||:|:|||||||||||||
Db          4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy         66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
              ||||||||||||||||||||||||||||||||||||:..|||:||||||||:|||||||||
Db         64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy        126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||.:----S 178
Db        124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy        186 QGDDRYICDRLYPDSLMWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
              :.|||||||:||:.|:||||||||||||||:|||||||||||||||||||||||||||||
Db        179 EADDRYICDRFYPNDLWVVFQFQHIMVGLLLPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy        246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
              ||||||||||||||||||:||||||||| :::||||.||:.||||||||||||||||||||
Db        239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy        306 PILYAFLGAKFKSSAQHALNSMSRGGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
              |||||||||||:.||||||.:.|||.|||||||||||||||||||||||||||||
Db        299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352


RESULT 10
US-10-225-567A-76
; Sequence 76, Application US/10225567A
; Publication No. US2003011379 8A1
;   GENERAL INFORMATION:
;   APPLICANT: LifeSpan Biosciences
;   APPLICANT: Brown, Joseph P.
;   APPLICANT: Burmer, Glenna C.
;   APPLICANT: Roush, Christine L.
;   TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;   FILE REFERENCE: 1920-4-4
;   CURRENT APPLICATION NUMBER: US/10/225,567A
;   CURRENT FILING DATE: 2001-12-19
;   PRIOR APPLICATION NUMBER: 60/257,144
;   PRIOR FILING DATE: 2000-12-19
;   NUMBER OF SEQ ID NOS: 2292
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 76
;   LENGTH: 352
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-225-567A-76
```

```
  Query Match             89.9%;  Score 1678.5;  DB 14; Length 352;
  Best Local Similarity   90.1%;  Pred. No. 1.1e-145;
  Matches  319;  Conservative   19;  Mismatches   11;  Indels    5;  Gaps     1;

Qy          6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM 65
              :|||||||||:||:||||||||:||||||:||.|.|||||||||:|:|||||||||||||
Db          4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy         66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
              ||||||||||||||||||||||||||||||||||||:..|||:||||||||:|||||||||
Db         64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy        126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||.:----S 178
Db        124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy        186 QGDDRYICDRLYPDSLMWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
              :.|||||||:||:.|:||||||||||||||:|||||||||||||||||||||||||||||
Db        179 EADDRYICDRFYPNDLWVVFQFQHIMVGLLLPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy        246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
              ||||||||||||||||||:||||||||| :::||||.||:.||||||||||||||||||||
Db        239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy        306 PILYAFLGAKFKSSAQHALNSMSRGGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
              |||||||||||:.||||||.:.|||.|||||||||||||||||||||||||||||
Db        299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352


RESULT 11
US-10-245-850-1
; Sequence 1, Application US/10245850
; Publication No. US2003012462 8A1
;   GENERAL INFORMATION:
;   APPLICANT: Burns, Jennifer M.
;   APPLICANT: Miao, Zhenhua
;   APPLICANT: Wei, Zheng
;   APPLICANT: Howard, Maureen C.
;   APPLICANT: Premack, Brett A.
;   APPLICANT: Schall, Thomas J.
;   APPLICANT: ChemoCentryx, Inc.
;   TITLE OF INVENTION: Compositions and Methods for Detecting and Treating
;   TITLE OF INVENTION: Diseases and Conditions Related to Chemokine Receptors
;   FILE REFERENCE: 019934-003310US
;   CURRENT APPLICATION NUMBER: US/10/245,850
;   CURRENT FILING DATE: 2002-09-16
;   PRIOR APPLICATION NUMBER: US 60/338,100
;   PRIOR FILING DATE: 2001-11-30
;   NUMBER OF SEQ ID NOS: 3
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 1
;   LENGTH: 352
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: CXCR4 chemokine receptor
US-10-245-850-1

  Query Match             89.9%;  Score 1678.5;  DB 14; Length 352;
  Best Local Similarity   90.1%;  Pred. No. 1.1e-145;
  Matches  319;  Conservative   19;  Mismatches   11;  Indels    5;  Gaps     1;

Qy          6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM 65
              :|||||||||:||:||||||||:||||||:||.|.|||||||||:|:|||||||||||||
Db          4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy         66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
              ||||||||||||||||||||||||||||||||||||:..|||:||||||||:|||||||||
Db         64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy        126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        124
```

US-10-021-660-123
; Sequence 123, Application US/10021660
; Publication No. US2003015292 6A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-123

Query Match   89.9%;   Score 1678.5; DB 14;   Length 352;
Best Local Similarity   90.1%;   Pred. No. 1.1e-145;
Matches 319; Conservative   19; Mismatches 11; Indels   5; Gaps   1;

Qy   6 VSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLLTGIVNGLVLVM 65
           ::|||||||||:|||:|:||| ||| :||:|||||||||:|||||||||||||||||
Db   4 ISIYTSDNYTEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVNGLVLVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
           |||||||||||||||||||||||||||:||||||| ||||:|:|||||:|||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFADVSQGDIS 185
           ||||||||||||||||||||||||||||||:|||||||||||||||:||
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
           ||||||||||||||||||||:|:|||||||:||||||||||||||||||||||||||||
Db   179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
           |||||||||||||||||:||||||||:||::||| |||:|||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
           |||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

RESULT 12
US-10-251-703-38
; Sequence 38, Application US/10251703
; Publication No. US2003014 8449A1
; GENERAL INFORMATION:
; APPLICANT: Kuliopulos, Athan
; APPLICANT: Covic, Lidija
; TITLE OF INVENTION: G Protein Coupled Receptor Agonists and Antagonists and
; TITLE OF INVENTION: Methods of Activating and Inhibiting G Protein Coupled
; TITLE OF INVENTION: Receptors Using the Same
; FILE REFERENCE: NEMC-215 CIP
; CURRENT APPLICATION NUMBER: US/10/251,703
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 09/841,091
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/198,993
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Pepducin
; OTHER INFORMATION: Peptide Sequence
US-10-251-703-38

Query Match   89.9%;   Score 1678.5; DB 14;   Length 352;
Best Local Similarity   90.1%;   Pred. No. 1.1e-145;
Matches 319; Conservative   19; Mismatches 11; Indels   5; Gaps   1;

Qy   6 VSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLLTGIVNGLVLVM 65
           ::|||||||||:|||:|:||| ||| :||:|||||||||:|||||||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVNGLVLVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
           |||||||||||||||||||||||||||:||||||| ||||:|:|||||:|||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFADVSQGDIS 185
           ||||||||||||||||||||||||||||||:|||||||||||||||:||
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
           ||||||||||||||||||||:|:|||||||:||||||||||||||||||||||||||||
Db   179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
           |||||||||||||||||:||||||||:||::||| |||:|||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
           |||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

RESULT 13

RESULT 14
US-10-014-322A-126
; Sequence 126, Application US/10014322A
; Publication No. US2003016 7129A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, Jr., John
; APPLICANT: Wilson, Carol
; APPLICANT: Tan Nehir, Christina
; APPLICANT: Kates, Steven
; APPLICANT: Krstenansky, John
; TITLE OF INVENTION: Binding Compounds and Methods for Identifying Binding Compounds
; FILE REFERENCE: CNS-008
; CURRENT APPLICATION NUMBER: US/10/014,322A
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/243,587
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 09/813,651
; PRIOR FILING DATE: 2001-03-20

Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S 178
Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
Db   179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238
Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298
Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

Query Match   89.9%;   Score 1678.5; DB 14; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels   5;   Gaps   1;

Qy   6 VSIYTSDNYSEEVGSGDYDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLPTIYFIIFLTGIVGNGLVILVM 65
             |||||||||.|:.||||||||||||||||: |||: || :|||||||||||||
Db   4 ISIYTSDNYTEMGSGDYDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
             ||||||||||||||||||||||||||||||||||:|:|: :||||||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGWIPALLLTIPDFIFADVSQGDIS 185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||-----S
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
             :.|||||: :|: |||||||||||||||||||||||||||||||||||||||||||||
Db   179 EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
             ||||||||||||||||||||||||||||||: :|| :||||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
             |||||||||||| :||||:||||||||||||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

Search completed: May 17, 2004, 22:38:46
Job time : 549 secs

Query Match   89.9%;   Score 1678.5; Pred. No. 1.1e-145;   DB 14; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels   5;   Gaps   1;

Qy   6 VSIYTSDNYSEEVGSGDYDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
             |||||||||.|:.||||||||||||||||: |||: || :||||||||||||||
Db   4 ISIYTSDNYTEMGSGDYDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
             ||||||||||||||||||||||||||||||||||||:|: :||||||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGWIPALLLTIPDFIFADVSQGDIS 185
             ||||||||||||||||||||||||||||||: :|||||||||||||||||-----S
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGWIPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
             :.|||||: :|:||:|||||||||||||||||||||||||||||||||||||||||||
Db   179 EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
             ||||||||||||||||||:|||||||||||: :||::|||||||||||||||||||||
Db   239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
             ||||||||||||:|||||:||||||||||||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 17, 2004, 21:33:47 ; Search time 44 Seconds
(without alignments)
784.835 Million cell updates/sec

Title: US-09-367-052-2
Perfect score: 1867
Sequence: 1 MEPISVSIYTSDNYSEEVGS.........KRGGHSVSTESESSSFHSS 359
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 9619526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1679.5	90.0	352	2	G00048	fusin (LESTRA) - c
2	1678.5	89.9	352	2	A45747	neuropeptide Y/pep
3	1631	87.4	353	2	S28787	neuropeptide Y/pep
4	585.5	31.4	367	2	JE0349	interferon-inducib
5	570.5	30.6	360	2	A57160	chemokine (C-C) re
6	551.5	29.5	355	2	JQ1231	interleukin-8 rece
7	550.5	29.5	378	2	A55735	G protein-coupled
8	547.5	29.3	356	2	S42096	interleukin-8 rece
9	544.5	29.2	358	2	A53752	interleukin-8 rece
10	544.5	29.2	360	2	A53611	interleukin-8 rece
11	537.5	28.8	378	2	B55735	lymphocyte-specifi
12	535	28.7	369	2	JC5068	G protein-coupled
13	533	28.5	355	2	JC4304	orphan G protein-c
14	530	28.4	360	2	JC4587	chemokine (C-C) re
15	526.5	28.2	327	2	S56162	MDCR15 protein - h
16	523.5	28.0	359	2	A48921	interleukin-8 rece
17	522.5	28.0	350	2	A39445	interleukin-8 rece
18	522.5	28.0	372	2	S26667	G protein-coupled
19	518.5	27.8	378	2	A45680	G protein-coupled
20	517.5	27.7	374	2	S42628	G protein-coupled
21	508.5	27.2	374	2	S32785	G protein-coupled
22	507	27.2	383	2	S55594	probable G protein
23	501	26.8	354	2	I58186	chemokine (C-C) re
24	496	26.6	374	2	I38450	G protein-coupled
25	494	26.5	355	2	JC5067	G protein-coupled
26	484	25.9	350	2	JN0621	chemokine (C-C) re
27	483.5	25.9	352	2	A43113	chemokine (C-C) re
28	479.5	25.7	360	2	JC2443	chemokine (C-C) re
29	476.5	25.5	355	2	G02436	chemokine (C-C) re
30	473.5	25.4	359	2	S15403	angiotensin II rec
31	472.5	25.3	355	2	A45177	chemokine (C-C) re
32	463.5	24.8	359	2	JC1104	angiotensin II rec
33	463.5	24.8	362	2	JN0694	angiotensin II rec
34	460	24.6	362	2	A30341	G protein-coupled
35	459.5	24.6	359	2	S44425	angiotensin II rec
36	459.5	24.6	359	2	A42656	angiotensin II rec
37	459.5	24.6	359	2	JH0621	angiotensin II rec
38	459.5	24.6	359	2	JC2134	angiotensin II rec
39	457.5	24.5	359	2	A48857	angiotensin II rec
40	453.5	24.3	359	2	JQ1516	angiotensin II rec
41	453	24.3	354	2	A23669	interleukin-8 rece
42	448	24.0	362	2	A39714	G protein-coupled
43	447.5	24.0	359	2	I49341	MIP-1 alpha recept
44	445.5	23.9	359	2	I39418	angiotensin II rec
45	442.5	23.7	359	2	JC1194	angiotensin II rec

ALIGNMENTS

RESULT 1
G00048
fusin (LESTRA) - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 26-Aug-1999
C;Accession: G00048
R;Tatsumi, M.
submitted to GenBank, July 1996
A;Reference number: H00048
A;Accession: G00048
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-352 <TAT>
A;Cross-references: GB:D86579; NID:g1468948; PID:g1468949
C;Superfamily: vertebrate rhodopsin

Query Match   90.0%; Score 1679.5; DB 2; Length 352;
Best Local Similarity   90.1%; Pred. No. 5.3e-140;
Matches 319; Conservative 17; Mismatches 13; Indels 5; Gaps 1;

```
Qy    6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM 65
          | ||||||||:|:|  ||||||||||||  ||||||:||||||||||||||||||||||
Db    4 ISIYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          |||||||||||||||||||||||||||||||||:|| |||||||||||||||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLYVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||:    S
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFASV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          | ||:|||| || || | |||||||||||||||||||||:|||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYVYVGISIDSFILLGVIKQGCDFBSIVHKWISTEALAFFHCCLN 305
          ||||||||||||||||||| :||||||||| ||:|||:|  | |||||||||| ||||||
Db  239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISTEALGFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          |||||||||||||:|||||| | |||||||||||||||||||||||||||||||
Db  299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 2
A45747
neuropeptide Y/peptide YY receptor Y3 - human
N;Alternate names: fusin; HM89; leukocyte-derived seven-transmembrane receptor LESTR; re
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-Nov-1999
C;Accession: A45747; A53103; I53306; I9444; I69203; S32761

R;Federsppiel, B.; Melhado, I.G.; Duncan, A.M.V.; Delaney, A.; Schappert, K.; Clark-Lewi
Genomics 16, 707-712, 1993
A;Title: Molecular cloning of the cDNA and chromosomal localization of the gene for a pu
A;Reference number: A45747; MUID:93315164; PMID:8325644
A;Accession: A45747
A;Molecule type: mRNA
A;Residues: 1-352 <FED>
A;Cross-references:GB:M99293; NID:g292516; PIDN:AAA16617.1; PID:g292517
R;Loetscher, M.; Geiser, T.; O'Reilly, T.; Zwahlen, R.; Baggiolini, M.; Moser, B.
J. Biol. Chem. 269, 232-237, 1994
A;Title: Cloning of a human seven-transmembrane domain receptor, LESTR, that is highly e
A;Reference number: A53103; MUID:94103215; PMID:8276799
A;Accession: A53103
A;Molecule type: mRNA
A;Residues: 1-352 <LOE>
A;Cross-references: EMBL:X71635; NID:g297099; PIDN:CAA50641.1; PID:g297100
R;Herzog, H.; Hort, Y.J.; Shine, J.; Selbie, L.A.
DNA Cell Biol. 12, 465-471, 1993
A;Title: Molecular cloning, characterization, and localization of the human homolog to t
A;Reference number: I53006; MUID:93319629; PMID:8329116
A;Accession: I53006
A;Molecule type: mRNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-352 <HER>
A;Cross-references: GB:L06797; NID:g414929; PIDN:AAA03209.1; PID:g414928
R;Jazin, E.E.; Yoo, H.; Blomqvist, A.G.; Yee, F.; Weng, G.; Walker, M.W.; Salon, J.; Laz
Regul. Pept. 47, 247-258, 1993
A;Title: A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its human homolo
A;Reference number: I59444; MUID:94052833; PMID:8234909
A;Accession: I59444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-352 <RE2>
A;Cross-references: GB:L01639; NID:g189313; PIDN:AAA16594.1; PID:g189314
R;Nomura, H.; Nielsen, B.W.; Matsushima, K.
Int. Immunol. 5, 1239-1249, 1993
A;Title: Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte chem
A;Reference number: I54751; MUID:94092629; PMID:7505609
A;Accession: I69203
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-352 <RES>
A;Cross-references: GB:D10924; NID:g219868; PIDN:BAA01722.1; PID:g219869
C;Genetics:
C;Gene: GDB:NPY3R; NPY3
A;Cross-references: GDB:230002; OMIM:162643
A;Map position: 2q21-2q21
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match   89.9%; Score 1678.5; DB 2; Length 352;
Best Local Similarity   90.1%; Pred. No. 6.5e-140;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy   6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVLVM 65
            ||||||||| ||||||||||||||||||||||||:||||||||||||||||||||||
Db   4 ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVLVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
           |||||||||||||||||||||||||||||||||:|:||||||:||||||||:|||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy   126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
           ||||||||||||||||||||||||||||||||::|||||||||||||||:|||||| |:|
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWLPALLLTIPDFIFANV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVIKQGCDFESIVHKWISITEALAFFHCCLN 245
           |||||||||:|:||||||||||||||||||||||||||:| ||:||||||||||||||||||:|:
Db   179 EADDRYICDRFYPNDLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
           |||||||||||||||||||||||||||||:||:| ||:|||||||||||||||||||:|:
Db   239 KTTVILILTFFACWLPYYIGISIDSFILLEIIQQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGGSLKILSKGKRGGHSSVSTESESSSFHSS 359
           ||||||||||||||||||||||||:||||||:|||||||:||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352

RESULT 3
S28787
neuropeptide Y/peptide YY receptor Y3 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-Aug-1999
C;Accession: S28787
R;Rimland, J.; Xin, W.; Sweetnam, P.; Saijoh, K.; Nestler, E.J.; Duman, R.S.
Mol. Pharmacol. 40, 869-875, 1991
A;Title: Sequence and expression of a neuropeptide Y receptor cDNA.
A;Reference number: S28787; MUID:9210053; PMID:1661837
A;Accession: S28787
A;Molecule type: mRNA
A;Residues: 1-353 <RIM>
A;Cross-references: EMBL:M86739
C;Superfamily: vertebrate rhodopsin
C;Keywords: appetite; G protein-coupled receptor; transmembrane protein

Query Match   87.4%; Score 1631; DB 2; Length 353;
Best Local Similarity   86.8%; Pred. No. 9.8e-136;
Matches 308; Conservative 26; Mismatches 15; Indels 6; Gaps 2;

Qy   6 VSIYTSDNYSE-EVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVLV 64
            :|||||||| ::||||||||||||||||||||||||||||||||||||||||||||
Db   4 IRIFTSDNYTEDDLGSGDYDSMKEPCFREENAHFNRIFLPTVYSIIFLTGIVGNGLVLV 63

Qy   65 MGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYS 124
           |||||||||||||||||||||||||||||||||:|:||||||:|||||||:||||||||
Db   64 MGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGKFLCKAVHVIYTVNLYS 123

Qy   125 SVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDI 184
           :||||||||||||||||||||||||||||||||::||||:|||||||||||||:-----DI 178
Db   124 SVLILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWLPAVLLTIPDLIFA-----DI 178

Qy   185 SQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKA 244
           :|:|||||||:|:|||||||||||||||||||||||||||||||||||:||||||:|||| 238
Db   179 KEVDERYICDRFYPSDLWLPYIVGLLLPGIVILSCYCIIISKLSHSKGYQKRKA 238

Qy   245 LKTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCL 304
           |||||||||||||||||||||||||||||:|||||:||:||||||||||||||||||||| 298
Db   239 LKTTVILILTFFACWLPYYIGISIDSFILLEIIQQGCEFRSTVHKWISITEALAFFHCCL 298

Qy   305 NPILYAFLGAKFKSSAQHALNSMSRGGSLKILSKGKRGGHSSVSTESESSSFHSS 359
           |||||||||||||||||||||||:|||||||||||||||:||||||||||||||| 353
Db   299 NPILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 353

RESULT 4
JE0349
interferon-inducible protein 10 (IP-10) receptor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jul-2000
C;Accession: JE0349
R;Tamaru, M.; Tominaga, Y.; Yatsunami, K.; Narumi, S.
Biochem. Biophys. Res. Commun. 251, 41-48, 1998
A;Title: Cloning of the murine interferon-inducible protein 10 (IP-10) receptor and its
A;Reference number: JE0349; MUID:99009219; PMID:9790904
A;Accession: JE0349
A;Molecule type: mRNA
A;Residues: 1-367 <TAM>
A;Cross-references: DDBJ:AB003174; NID:g3798731; PIDN:BAA34045.1; PID:g3798732
C;Comment: This protein is important for lymphocyte trafficking to lymphoid organs.
C;Superfamily: vertebrate rhodopsin

Query Match   31.4%; Score 585.5; DB 2; Length 367;
Best Local Similarity   35.4%; Pred. No. 8.2e-44;
Matches 123; Conservative 69; Mismatches 130; Indels 25; Gaps 7;

```
Qy    17 EVGSGDYD-SNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGYQKLRSMT  75
         |   | :  :|:::::|| |||||| |||||||||||||||| ||:|||||||:|:||||
Db    28 DYGENESDFSDSPPCPQDFSLNFDRTFLPALYSLLFLLGLLGNGAVAAVLLSQRTALSST  87

Qy    76 DKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVLILAFISLD 135
         | :||||:|||:||::||| ||||||  |::||:|||||  :: ||:|||| ||||:| |
Db    88 DTFLLHLAVADVLLVLTLPLWAVDAAVQWVFGPGLCKVAGALFNINFYAGAFLLACISPD 147

Qy   136 RYLAIVHATNSQR--PRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQGDDRYI- 192
         ||| |||||:  :  ||: ::|:| : :::  ::: ::|::|||:      |:|: |::
Db   148 RYLSIVHATQIYRRDPRVRVALTCIV--VWGLCLLFALFPDFIY------LSANVDQRLN 198

Qy   193 ---CDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTV 249
            |:|      : |  :::: || :| |::||:: :|  ||  | |: |:|:||: |||
Db   199 ATHCQYNFP-QVGRTALRVLQLVAGFLLPLLVMAYCYAHILAVLLVSRGQRRFRAMRLVV 257

Qy   250 ILILAFPACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPILY 309
         : : |  | ||:|::|::|     :|::   |  ::| :: :: |  | |::||||| ||
Db   258 VVVAARAFAVCWTPYHLVVLVDILMDVGVLARNCGRESHVDVAKSVTSGMGYMHCCLNPLLY 317

Qy   310 AFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSF 356
         ||:| ||:: :                 :|:|:        :|  |
Db   318 AFVGVKFREQWMMLFTRLGRSD------QRGPQRQPSSSRRESSW 356
```

RESULT 5
A57160
chemokine (C-C) receptor 4 - human
N;Alternate names: C-C CKR-4
C;Species: Homo sapiens (man)
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C;Accession: A57160
R;Power, C.A.; Meyer, A.; Nemeth, K.; Bacon, K.B.; Hoogewerf, A.J.; Proudfoot, A.E.I.; W
J. Biol. Chem. 270, 19495-19500, 1995
A;Title: Molecular cloning and functional expression of a novel CC chemokine receptor cD
A;Reference number: A57160; MUID:95370289; PMID:7642634
A;Accession: A57160
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-360 <POW>
A;Cross-references: GB:X85740; NID:g1370103; PIDN:CAA59743.1; PID:g971452
A;Note: source clone KS-5
C;Genetics:
A;Gene: GDB:CMKBR4
A;Cross-references: GDB:677463
A;Map position: 3p21-3p21
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F;40-65/Domain: transmembrane #status predicted <TM1>
F;76-97/Domain: transmembrane #status predicted <TM2>
F;112-133/Domain: transmembrane #status predicted <TM3>
F;151-175/Domain: transmembrane #status predicted <TM4>
F;208-226/Domain: transmembrane #status predicted <TM5>
F;243-264/Domain: transmembrane #status predicted <TM6>
F;291-308/Domain: transmembrane #status predicted <TM7>
F;29-276,110-187/Disulfide bonds: #status predicted
F;72,350/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicte
F;145/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;183,194/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match   30.6%;   Score 570.5;   DB 2;   Length 360;
Best Local Similarity   37.4%;   Pred. No. 1.7e-42;
Matches 120;   Conservative 63;   Mismatches 125;   Indels 13;   Gaps 6;

```
Qy     1 MEPISVSIYTSDNYSEVGSGD--YDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGN  58
         | | |::||:| :  || ||   |::  |||  : :   || ||| |:|::|:  ::||
Db     1 MNPTDIADTTLD---ESIYSNYILYESIPKPCTKEGIKAFGELFLPPLYSLVFVFGLLGN  57

Qy    59 GLVILVMGYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIY 118
         | |:||:  :|||||||:|  :|:|:|||||:::|||   :|  |||| ||||:: : |
Db    58 SVVVLVLFKYKRLRSMTDVYLLNLAISDLLFVFSLPFWGYYAADQWVFGLGLCKMISWNY 117
```

```
Qy   119 TVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFAD 178
          | :||| :::: : |:|||||||| : :|:  : : : ::  : :: : : |:| :| :
Db   118 LVGFYSGIFFVMLMSIDRYLAIVHAVFSLRARTITYGVITSLATWSVAVFASLPGFLFST 177

Qy   179 VSQGDISQGDDRYICDRLYP--DSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHS 236
              :  : :   ::|    |  :: : : :      :: :| ||::| |||:: ::| |
Db   178 C----YTERNHTY-CKTKYSLNSTTWKVLSSLEINIGLVIPLGIMLFCYSMIIRTLQHC 232

Qy   237 KGHQKRKALKTVTVLILAFPACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEA 296
         | :::: | | :  :::|| | | |||:: : ||   |:   :||:||  : : : : | 
Db   233 KNEKKNKAVKMIFAVVVLFLGFWTPYNIVLFLETLVELEVL-QDCTFERYLDYAIQAET 291

Qy   297 LAFFHCCLNPILYAFLGAKFK 317
         |||.|||||| :| | | ||
Db   292 LAFVHCCLNPIIYFFLGEKFR 312
```

RESULT 6
JQ1231
interleukin-8 receptor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Nov-1999
C;Accession: JQ1231; A46483
R;Beckmann, M.P.; Munger, W.E.; Kozlosky, C.; VandenBos, T.; Price, V.; Lyman, S.; Gerar
Biochem. Biophys. Res. Commun. 179, 784-789, 1991
A;Title: Molecular characterization of the interleukin-8 receptor.
A;Reference number: JQ1231; MUID:91378994; PMID:1898400
A;Accession: JQ1231
A;Molecule type: DNA
A;Residues: 1-355 <BEC>
A;Cross-references: GB:M74240; NID:g165438; PIDN:AAA31375.1; PID:g165439
R;Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
J. Immunol. 148, 1261-1264, 1992
A;Title: Characterization of complementary DNA clones encoding the rabbit IL-8 receptor.
A;Reference number: A46483; MUID:92148149; PMID:1737938
A;Accession: A46483
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-355 <LEE>
A;Cross-references: GB:M82873; NID:g165440; PIDN:AAA31376.1; PID:g165441
A;Experimental source: neutrophils
A;Note: sequence extracted from NCBI backbone (NCBIN:81526, NCBIP:81530)
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match   29.5%;   Score 551.5;   DB 2;   Length 355;
Best Local Similarity   35.0%;   Pred. No. 7.7e-41;
Matches 121;   Conservative 72;   Mismatches 142;   Indels 11;   Gaps 6;

```
Qy    12 DNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGYQKKL  71
         |::  :.|:  :: :  || :: :  ::| : |:::: ::: :: ||.||::|:: |::
Db    17 DEFANATGMPPVEKDYSPCLVVTQT-LNKYVVVVIYAIVFLLSLLGNSLVMLVILYSRSN  75

Qy    72 RSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVLILAF 131
         |:||| | | |::|||||::::|||| :: :|| ||  |||.: :: ::| :||::|||
Db    76 RSVTDVYLLNLAMADLLFALTMPIWAVSKEKGWIFGTPLCKVVSLVKEVNFYSGILLLAC 135

Qy   132 ISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQGGDRY 191
         ||:|||||||||   : :::| : || ::|:|  : |:: || ||:       ||   |
Db   136 ISVDRYLAIVHATRTLTQKRHLV-KFICLGIWALSLILSLPFFLFRQV---FSPNNSSP 190

Qy   192 IC--DRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTV 249
         :|  : |   :: : :::: : : |::||::|:: || : : ::: | :: :|:| ::
Db   191 VCYEDLGHNTAKWRMVLRILPHTFGFILPLLVMLFCYGFTLRTLFQAHMGQKHRAMRVIF 250

Qy   250 ILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPILY 309
         : :: || | ||| :::  ::   :|:   : : .: : : :    | | ||||||: |
Db   251 AVLIFLLCWLPYNLVLLADTLMRTHVIQETCQRRNDIDRALDATEIIGFLHSCLNPIIY 310

Qy   310 AFLGAKFKSSAQHALNSMSRG-SSLKILSKGKRGGHSSVSTESESS 354
         ||:| :|:.     |: ::|  :|||       :|::: | :::
Db   311 AFIGQNFRNGFLKML--AARGLISKEFLTRHRVTSYTSSTNVPSN 354
```

```
RESULT 7
A55735
G protein-coupled receptor EBI1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 24-Nov-1999
C;Accession: A55735
R;Schweickart, V.L.; Raport, C.J.; Godiska, R.; Byers, M.G.; Eddy Jr., R.L.; Shows, T.B.
Genomics 23, 643-650, 1994
A;Title: Cloning of human and mouse EBI1, a lymphoid-specific G-protein-coupled receptor
A;Reference number: A55735; MUID:95154835; PMID:7851893
A;Accession: A55735
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-378 <SCH>
A;Cross-references: GB:L31580; NID:g468340; PIDN:AAA74232.1; PID:g468341
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor

Query Match     29.5%;  Score 550.5;  DB 2;  Length 378;
Best Local Similarity   33.2%;  Pred. No. 1e-40;
Matches 117;  Conservative  78;  Mismatches 150;  Indels  7;  Gaps  3;

Qy   11 SDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVNGLVILVMGYQKK 70
            :|:||: |||:| :|||:|||||||||||||||||:: :: :||
Db   29 TDDYIGENTTVDYTLYESVCFKKDVRNFKAWPLPLMYSVICFVGLLGNGLVILTYIYFKR 88

Qy   71 LRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHIIYTVNLYSSVLILA 130
            |:||:||||||| |:||:|||| ||||  |||||||||||||||  :||||:|:
Db   89 LKTMTDTYLLNLAVADILFLLILPFWAYSEAKSWIFGVYLCKGIFGIYKLSFFSGMLLLL 148

Qy  131 FISLDRYLAIVHAT--NSQRPRKLLAEKAVYYGVWIPALLLTIPDFIFADVSQGDISQGD 188
            ||| |||||:||:    |:|:|  |:|||  | ||:||:|||||:  |:|||:|
Db  149 CISIDRYVAIVQAVSRHRHRARVLLISKLSCVGIWMLALFLSIPELLYSGLQK---NSGE 205

Qy  189 DRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTT 248
            |||||||||||  ::|: :||: ||:||||||:|   ||:::
Db  206 DTLRCSLVSAQVEALITQVAQMVFGFLVPMLAMSFCVLIIIRTLLQARNFERNKAIKVI 265

Qy  249 VILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPIL 308
            ::|||| |||  || :| |||  ::|:||
Db  266 IAVVVVFIVFQLPYNGVVLAQTVANFNITNSSCETSKQLNIAYDVTYSLASVRCCVNPFL 325

Qy  309 YAFLGAKFKSSAQHALNSMSRGSS--LKILSKGKRGGHSSVSTESESSSFHS 358
            |||  |||||||||||||||||||  |||| |:
Db  326 YAFIGVKFRSDLFKDLGCLSQERLRHWSSCRHVRNASVSMBAETTTTFS 377
```

```
RESULT 8
S42096
interleukin-8 receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
C;Accession: S42096
R;Gobl, A.E.; Wang, S.; Zhou, Y.; Oeberg, K.
submitted to the EMBL Data Library, February 1994
A;Description: Molecular cloning of the rat IL8 receptor.
A;Reference number: S42096
A;Accession: S42096
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-356 <GOB>
A;Cross-references: EMBL:X77797
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match     29.3%;  Score 547.5;  DB 2;  Length 356;
Best Local Similarity   34.8%;  Pred. No. 1.7e-40;
Matches 129;  Conservative  67;  Mismatches 124;  Indels  51;  Gaps 11;

Qy   12 DNYS-EEVGSGDYD---------SNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGN 58
            ||:| ||||| :|         ::|||||||||||||||||||||||||| ::|||
Db    7 DNFSLEDFFSGDIDSYNFSSDPPTLSDAAPC-PSANLDINRXAVVVIYVLVTLLSLVGN 65

Qy   59 GLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHIIY 118
            :|::|: ||:| :|||:|||||||| |:||:|||| |||| ||||||||||||| :|||
Db   66 SLVMLIVLYNRSTCGSVTDVVLNLAIADLLFALTLPVWAASKVNGWIFGSFLCKVFSFLQ 125

Qy  119 TVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYYGVWIPALLLTIPDFIFAD 178
            |: :|||:|:|| ||:|||||||||:: ||:|:|  |: :   |: :|:|:|| |:|:
Db  126 BITFYSSVLLLACISMDRYLAIVHATSTLIQKRHLV-KFVCITMWFLSLVLSLPIFIL-- 182

Qy  179 VSQGDISQGDDRYICDRLYPD-----SLWMVVFQFQHIMVGLILPGIVILSCYCIIISKL 233
               ||:|    |::  :|:|     ::||:| | |: ||||: |||:|:|:||   ::|
Db  183 --RTTVKANPSTVVC---YENIGNNTSKWRVLRILPQTYGFLLPLLIMLFCYGFTLRTL 237

Qy  234 SHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISI 293
            | :  :|:|:|: |:|:|:|||  ||||| :| || ::|:| ||  |    ||| ::
Db  238 FKAHMGQKHRAMRVIFAVVLFLLCWLPYNIVLFTTDTLMRTKLIKETCERQNEINK---A 294

Qy  294 TEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSK---------GKRGGH 344
             : || ||||||::|||:| ||||     ||:      ||:|::         |:
Db  295 SBILGFLHSCLNPIIYAFIGQKFR------HGL--------LKIMANYGLVSKEFLAKEGRP 342

Qy  345 SSVSTESESSS 355
             ||::   :::
Db  343 SFVGSSSANTS 353
```

```
RESULT 9
A53752
interleukin-8 receptor (clone 5B1a) - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: A53752
R;Prado, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navarro
J.
J. Biol. Chem. 269, 12391-12394, 1994
A;Title: Molecular characterization of a novel rabbit interleukin-8 receptor isotype.
A;Reference number: A53752; MUID:94230294; PMID:8175642
A;Accession: A53752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-358 <PRA>
A;Cross-references: GB:L24445; NID:g437661; PIDN:AAA31378.1; PID:g437662
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match     29.2%;  Score 544.5;  DB 2;  Length 358;
Best Local Similarity   34.7%;  Pred. No. 3.2e-40;
Matches 128;  Conservative  70;  Mismatches 132;  Indels  39;  Gaps  9;

Qy    9 YTSDNYSEEVGSGDYDS----------NKEPCFRDENVHFNRIFLPTIYFIIFLTGIV 56
           :| ||:||||:| :|           ||||||||||||||||||||||||| :||
Db    4 FTWENYSYEDFFGDFSNYSYSTDLPPTLLDDSAPC-RSESLETNSYVVLITYILVFLLSLL 62

Qy   57 GNGLVLIVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHI 116
           ||:||::|: || :|||:||| |:||||||||| |||:||||  |||| ||:|||| :
Db   63 GNSLVMLVILYSRSTCSVTDVYLLNLAIADLLFATTLPIWAASKVHGWTFGTPLCKVVSL 122

Qy  117 IYTVNLYSSVLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYYGVWIPALLLTIPDFIF 176
            :: |:|| :||||| |:||||||||| : |:|:| |:  :   |  :|:|:|| |::
Db  123 VKEVNFYSGILLLACISVDRYLAIVHATRTMIQKRHLV-KFICLSMWGVSLILSLPILLF 181

Qy  177 ADVSQGDISQGDDRYICDRLYPD-----SLWMVVFQFQHIMVGLILPGIVILSLSCYCIIS 231
              :  :     |::  :|:|     ::|:| : |: ||| | |||:|::| ||  ::
Db  182 RNA-----IFPPNSSPVC---YEDMGNSTAKWRMVLRILPQTFGFILPLLVMLFCVVFTLR 234

Qy  232 KLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWI 291
            | :  :|:|:| : ||:|:|:|||  |||| :| || ::|:|:||:  |    ||  :
Db  235 TLFQAHMGQKHRAMRVIFAVVLIFLLCWLPYNLVLLTDTLMRTHVIQETCERRNDIDRAL 294

Qy  292 SITEALAFFHCCLNPILYAFLGAKF----KSSAQHALNSMSRGSSLKILSKGKRGGHSS 346
             : : || ||||||::|||:| ||:    |::|:      ||:|:|||: |:
Db  295 DATBILGFLHSCLNPIIYAFIGQKFRYGLLKILAANGL------ISKEFLAKESR-PSF 346

Qy  347 VSTESEBSS 355
```

Db   239 FKAHMGQKHRAMRVIFAVVLIFLLCWLPYNLVLLADTLMRTQVIQETCERRNHIDRALDA 298

Qy   294 TEALAPFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTES-- 351
          || || ||||||| |:|||::|:
Db   299 TEILGILHSCLNPLIYAFIGQKFR----HGL------LKILAI----HGLISKDSLP 341

Qy   352 -------ESSSFHSS 359
              :|| |||
Db   342 KDSRPSFVGSSSGHTS 357

RESULT 11
B55735
lymphocyte-specific G protein-coupled receptor EBI1 - human
N;Alternate names: Burkitt's lymphoma receptor 2; Epstein-Barr virus induced protein 1
C;Species: Homo sapiens (man)
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 19-May-2000
C;Accession: B55735; S52443
R;Schweickart, V.L.; Raport, C.J.; Godiska, R.; Byers, M.G.; Eddy Jr., R.L.; Shows, T.B.
Genomics 23, 643-650, 1994
A;Title: Cloning of human and mouse EBI1, a lymphoid-specific G-protein-coupled receptor
A;Reference number: A55735; MUID:95154835; PMID:7851893
A;Accession: B55735
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-378 <SCH>
A;Cross-references: GB:L31581; NID:g468319; PIDN:AAA74231.1; PID:g468320
R;Burgstahler, R.; Kempkes, B.; Staeube, K.; Lipp, M.
submitted to the EMBL Data Library, February 1995
A;Description: The expression of the chemokine receptor BLR2/EBI1 is specifically transa
A;Reference number: S52443
A;Accession: S52443
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 21-378 <BUR>
A;Cross-references: EMBL:X84702
C;Genetics:
A;Gene: GDB:CMKBR7; EBI1; BLR2; CCR7
A;Cross-references: GDB:342065; OMIM:600242
A;Map position: 17q12-17q21.2
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor

Query Match   28.8%;   Score 537.5; DB 2;   Length 378;
Best Local Similarity 32.1%;   Pred. No. 1.4e-39;
Matches 113;   Conservative 79;   Mismatches 153;   Indels   7;   Gaps   3;

Qy   11 SDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVMGYQKK 70
          ||||: | ::|| :|| | ||||||:|| ||||:|||||||||||:|||:||:|||||||
Db   29 TDDYIGDNTTVDYTLFESLCSKKDVRNFKAWFLPIMYSIICFVGLLGNGLVVLTYIYFKR 88

Qy   71 LRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHIIYTVNLYSSVLILA 130
          | ||||||||||||||||| :::|| | :: :|||||||||||||| |||||| |:||||
Db   89 LKTMTDTYLLNLAVADILFLLTLPFWAYSAAKSWVFGVHFCKLIFAIYKMSFFSGMLLLL 148

Qy   131 FISLDRYLAIVHATNS--QRPRKLLAEKAVYYGVWIPALLLTIPDFIFADVSQGDISQGD 188
          |:|| ||| : :||: |::|||::| :: :| ::||| ||: || :::::|
Db   149 CISIDRYVAIVQAVSAHRRARVLLISKLSCVGIWILATVLSIPELLYSDLQR--SSSE 205

Qy   189 DRYICDRLYPDSLMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTT 248
          |||| ||| ::||:::||||| ||| | |||||| | :: ||||||||||:|:|:
Db   206 QAMRCSLITEHVEAFITIQVAQMVIGFLVPLLAMSFCYLVIIRTLLQARNFERNKAIKVI 265

Qy   249 VILILAFFACWLPYYVGISDSFILLGVIKQGCDPESIVHKWISITEALAFFHCCLNPIL 308
          : :|:|||| ||| || |:|| ::|: :||||: |
Db   266 IAVVVFVIFQLPYNGVVLAQTVANFNITSSTCELSKQLNIAYDVTYSLACVRCCVNFFL 325

Qy   309 YAFLGAKFKSSAQHALNSMS--RGSSLKILSKGKRGGGHSSVSTESESSSFHS 358
          |||| ||| | |:|:|:|: |: : :|||:|:
Db   326 YAFIGVKFRNDLFLFKDLGCLSGEQLRQWSSCRHIRRSMSVEAETTTTFS 377

RESULT 12

```
JC5068
G protein-coupled receptor CKR-L3 - human
C;Species: Homo sapiens (man)
C;Date: 31-Jan-1997 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C;Accession: JC5068
R;Zaballos, A.; Varona, R.; Gutierrez, J.; Lind, P.; Marquez, G.
Biochem. Biophys. Res. Commun. 227, 846-853, 1996
A;Title: Molecular cloning and RNA expression of two new human chemokine receptor-like g
A;Reference number: JC5067; MUID:97040707; PMID:8886020
A;Accession: JC5068
A;Molecule type: DNA
A;Residues: 1-369 <ZAB>
A;Cross-references: EMBL:Z79784; NID:g1668737; PIDN:CAB02144.1; PID:g1668738
C;Comment: This protein belongs to the family of alpha chemokine receptors.
C;Genetics:
A;Gene: GDB:CMKBR6; STRL22; GPR29; CCR6; CKR-L3; GPR-CY4
A;Cross-references: GDB:5370639; OMIM:601835
A;Map position: 6q27-6q27
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein
F;42-68/Domain: transmembrane #status predicted <TM1>
F;79-99/Domain: transmembrane #status predicted <TM2>
F;115-136/Domain: transmembrane #status predicted <TM3>
F;160-180/Domain: transmembrane #status predicted <TM4>
F;212-233/Domain: transmembrane #status predicted <TM5>
F;250-271/Domain: transmembrane #status predicted <TM6>
F;292-315/Domain: transmembrane #status predicted <TM7>


Query Match     28.7%;  Score 535;  DB 2;  Length 369;
Best Local Similarity   33.9%;  Pred. No. 2.3e-39;
Matches 124;  Conservative 72;  Mismatches 136;  Indels 34;  Gaps 9;

Qy    10  TSDNYSEEVGSGDY--DSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGY   67
              ::::|| |: ::|   ||||||||||||||||||||||  | :||| || ||:::
Db     9  SSEDYFVSVNTSYYSVDSEMLLCSLQEVRQFSRLFVPIAYSLICVFGLLGNILVVTFAF   68

Qy    68  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVD-AWADWYFGKFLCKAVHIIYTVNLYSSV  126
              ||:|||||||||||||||||||||||||||   :|||||||||||||||||||||:|
Db    69  YKKARSMTDVYLLNMAIADILFVITLPFWAVSHATGAWVFSNATCKLLKGIYAINFNCGM  128

Qy   127  LILAFISLDRYLAIVHATNS-QRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDI   184
              |:|| || |||||||||||  ||||||||||: :::|||||||||||  ||||||
Db   129  LLLTCISMDRYIAIVQATKSFRLRSRTLPRSKIICLVVWGLSVIISSTFVF---NQKYN  185

Qy   185  SQGDDRYICDRLY----PDSLMWVVFQPQHIMVGLILPGIVILSCYCIIISKLSHSKGHQ  240
              :| |:  ||| |    ||||||||||||||||||||||||||||||||||::|:|
Db   186  TQGSD--VCEPKYQTVSEPIRWKLLMLGLELLFGFFIPLMFMFCYTFIVKTLVQAGNSK   243

Qy   241  KRKALKTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFF  300
              | ||  |||||||||||||||||||||||||||||| ||:||||||||||||||||
Db   244  RHKAIRVIIAVVLVFLACQIPHNMVLVTA-ANLGKMNRSCQSEKLIGYTKTVTEVLAFL  302

Qy   301  HCCLNPIYAFLGAKFKSSAQHALNSMRGSSLKILS------KGKRGGHSSVSTESES    353
              ||||||||||:||||||
Db   303  HCCLNPVLYAFIGQKF--------------RNYFIKILKDLWCVRRKYKSSGFSCAGRYSEN  350

Qy   354  SSFHSS  359
              :|
Db   351  ISRQTS  356


RESULT 13
JC4304
orphan G protein-coupled receptor - human
N;Alternate names: V28 protein
C;Species: Homo sapiens (man)
C;Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 19-May-2000
C;Accession: JC4304
R;Raport, C.J.; Schweickart, V.L.; Eddy Jr., R.L.; Shows, T.B.; Gray, P.W.
Gene 163, 295-299, 1995
A;Title: The orphan G-protein-coupled receptor-encoding gene V28 is closely related to g
A;Reference number: JC4304; MUID:96011651; PMID:7590284
A;Accession: JC4304
A;Molecule type: mRNA
A;Residues: 1-355 <RAP>
A;Cross-references: GB:U20350; NID:g665580; PIDN:AAA91783.1; PID:g665581
A;Experimental source: peripheral blood mononuclear cell
C;Comment: This protein is a cell-surface receptor which recognizes extracellular signal
C;Comment: This protein is a key regulator of many immune and homeostatic responses, and
C;Genetics:
A;Gene: v28
A;Map position: 3pter-p21
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; lymphokine; transmembrane protein
F;35-57/Domain: transmembrane #status predicted <TM1>
F;66-88/Domain: transmembrane #status predicted <TM2>
F;104-125/Domain: transmembrane #status predicted <TM3>
F;146-165/Domain: transmembrane #status predicted <TM4>
F;197-217/Domain: transmembrane #status predicted <TM5>
F;230-254/Domain: transmembrane #status predicted <TM6>
F;275-296/Domain: transmembrane #status predicted <TM7>


Query Match     28.5%;  Score 533;  DB 2;  Length 355;
Best Local Similarity   33.1%;  Pred. No. 3.2e-39;
Matches 120;  Conservative 65;  Mismatches 146;  Indels 32;  Gaps 7;

Qy    12  DNYSEEVGSG-DYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGVGNGLVILVMGYQKK   70
              ||||||||| ||  ::|: |||||||||||||||||| || ::||| ||::|:
Db     2  DQFPESVTENFEYDDLAEACYIGDIVVFGTVFLSIFYSVIFAIGLVGNLLVVFALTNSKK  61

Qy    71  LRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVLILA  130
              |:||||||||||||||||||||||   :||||||||||||:|||||||||:||| |
Db    62  PKSVTDIYLLNLALSDLLFVATLPFWTHYLINEKGLHNAMCKFTTAFFIGFFGSIFFIT  121

Qy   131  FISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQGDDR  190
              |||||||||:|||||  ||||||||||| |: |||||||||||  ||||||||   :|
Db   122  VISIDRYLAIVLAANSMNNRTVQHGVTISLGVWAAAILVAAPQFMFTKQKENE--------  174

Qy   191  YICDRLYPD--SLMWVVFQPQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT   247
              ||| | |  ||:|| || |||||||||||||||||||  |:|| || ||::| :|
Db   175  -CLGDYPEVLQEIWPVLRNVETNFLGFLLPLLIMSYCYRIIQTLFSCKNHKKAKAIKL   232

Qy   248  TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI  307
              |:| |||| |||||||||||||| :||||||||| || ||||||||||||||||:||||:|
Db   233  ILLVIVFFLFWTPYNVMIFLETLKLYDFF-PSCDMRKDLRLALSVTETVAFSHCCLNPL   291

Qy   308  LYAFLGAKFKSSAQHALNSMRGSSLKIL-------SKGKRGGHSSVSTESESSSF      356
              ||||| |||:|| :  :| ||| |:|         :|| || : : |||| |||
Db   292  IYAFAGEKFRRYLYHLY-----GKCLAVLCGRSVHVDFSSSESQRSRHGSVL--SSNFTY  344

Qy   357  HSS  359
              |:|
Db   345  HTS  347


RESULT 14
JC4587
chemokine (C-C) receptor 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Mar-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000
C;Accession: JC4587
R;Hoogewerf, A.J.; Black, D.; Proudfoot, A.E.I.; Wells, T.N.C.; Power, C.A.
Biochem. Biophys. Res. Commun. 218, 337-343, 1996
A;Title: Molecular cloning of murine CC CKR-4 and high affinity binding of chemokines to
A;Reference number: JC4587; MUID:96136324; PMID:8573157
A;Accession: JC4587
A;Molecule type: mRNA
A;Residues: 1-360 <HOO>
A;Cross-references: EMBL:X90862; NID:g1167851; PIDN:CAA62372.1; PID:g1167852
A;Experimental source: thymus
C;Genetics:
A;Gene: cc ckr-4
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; phosphoprotein; receptor; thymus
F;2,183,194/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;72,202,350/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predi
```

Search completed: May 17, 2004, 22:19:54
Job time : 45 secs

F;145/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

```
Query Match           28.4%;    Score 530;  DB 2;   Length 360;
Best Local Similarity 36.4%;     Pred. No. 6e-39;
Matches 108;  Conservative   62;  Mismatches 119;  Indels   8;  Gaps   4;

Qy   23  YDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHL   82
             |::  :  |:::  || ||| :|| :|:|:  ::|:| |:|:: |:|:
Db   22  YESMPKPCTKEGIKAFGEVFLPPIYSLVFLLGLFGNSVVVLVLFKYKRLKSMTDVYLLNL   81

Qy   83  SVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVH  142
             :::||||| |||| | | |  | || |:|||||  :|||||||||||||||||||
Db   82  AISDLLFVLSLPFWGYYAADQWVFGLGLCKIVSMYLVGFYSGIFFIMLMSIDRYLAIVH  141

Qy  143  ATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYP--DS  200
             : :   |    :::   :|::   : :   :  :|:: :|   :
Db  142  AVFSLKARTLTYGVITSLITWSVAVFASLPGLLFSTC---YTEHHTY-CKTQYSVNST  196

Qy  201  LWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTVILILAFFACWL   260
             |  ::| |::|:| || || |:  :| : |:  :: | |:|:|| :|::|||:
Db  197  TWKVLSSLEINVLGLLIPLGIMLFWYSMIIRTLQHCKNEKKNRAVRMIFGVVVLFLGFWT  256

Qy  261  PYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFK   317
             |  :| || : |  ||| |:| ||  ||:: :  |:| || :| ||:||||
Db  257  PYNVLFLETLVELEVL-QDCTLERYLDYAIQATETLGFIHCCLNPVIYFFLGEKFR   312
```

RESULT 15
S56162
MDCR15 protein - human
C;Species: Homo sapiens (man)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999
C;Accession: S56162
R;Barella, L.; Loetscher, M.; Tobler, A.; Baggiolini, M.; Moser, B.
Biochem. J. 309, 773-779, 1995
A;Title: Sequence variation of a novel heptahelical leucocyte receptor through alternati
A;Reference number: S56162; MUID:95366951; PMID:7639692
A;Accession: S56162
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-327 <BAR>
A;Cross-references: EMBL:X68829; NID:g840783; NID:g840784
C;Superfamily: vertebrate rhodopsin

```
Query Match           28.2%;    Score 526.5;  DB 2;   Length 327;
Best Local Similarity 34.6%;     Pred. No. 1.1e-38;
Matches 111;  Conservative   66;  Mismatches 139;  Indels   5;  Gaps   4;

Qy   38  FNRIFLPTIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWA   97
             :|:|   ||| || ::|| ||:|::  :::: ::       ||||||| :  |||:
Db    4  FKAVFVPVAYSLIFLLGVIGNVLVILERHRQTRSSTETFLFHLAVADLLLVFILDPFAV   63

Qy   98  VDAMADWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHAINSQRPRKLLAEKA  157
             :  | |::   | :: |:| || |  :|::|| | |||||| :: |:
Db   64  AEGSVGWVLGTFLCKTVIALHKVNFYCSSLLLACIAVDRYLAIVHAVHAYRHRRLLSIHI  123

Qy  158  VYVGVWIPALLLTIPDFIFADVSQGDISQGDDR--YICDRLYPDSLWMVVFQFQHIMVGL  215
             :|  |  | ||: :      || : :|:     :|| || : || || |||: |
Db  124  TCGTIWLVGFLLALPEILFAKVSQGHHNNSLPRCTFSQENQAETHAWF-TSRFLYHVAGF  182

Qy  216  ILPGIVILSCYCIIISKLSHSKGH-QKRKALKTVILILAFFACWLPYYVGISIDSFLLL   274
             |::| :: ::::     :   |::    :::::|: ||||| |||: :::| :|:  :
Db  183  LLPMLVMGWCYVGVVHRLRQAQRRPQRQKAVRVAILVTSIFFLCWSPYHIVIFLDTLARL  242

Qy  275  GVIKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK  334
             |::: | |     : | | || |:   || || ||| :|: |
Db  243  KAVDNTCKLNGSLPVAITMCEFLGLAHCCLNPMLYTFAGVKFRSDLSRLLTKLGCTSPAS  302

Qy  335  ILSKGKRGGHSSVTESESSS  355
             :|: :     ||| ::::
Db  303  LCQLFPSWRRSSLS-ESENAT  322
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     May 17, 2004, 16:16:00 ; Search time 27 Seconds
                                 (without alignments)
                                 692.340 Million cell updates/sec

Title:    US-09-367-052-2
Perfect score:  1867
Sequence:  1 MEPISVSIYTSDNYSEEVGS.........KRGGHSSVSTESESSSFHSS 359

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:  141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:     141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SwissProt_42:*

  Pred. No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1867	100.0	359	1	CCR4_MOUSE	P70658 m c-x-c che
2	1744.5	93.4	349	1	CCR4_RAT	O08565 rattus norv
3	1686.5	90.3	352	1	CCR4_PAPAN	P56491 papio anubi
4	1679	90.0	352	1	CCR4_MACFA	Q28474 macaca fasc
5	1678.5	89.9	352	1	CCR4_HUMAN	P30991 homo sapien
6	1676.5	89.8	352	1	CCR4_MACMU	P79394 macaca mula
7	1668.5	89.4	352	1	CCR4_CERTO	O62747 cercocebus
8	1643	88.0	353	1	CCR4_FELCA	P56498 felis silve
9	1631	87.4	353	1	CCR4_BOVIN	P25590 bos taurus
10	885.5	47.4	192	1	CCR4_SHEEP	Q28553 ovis aries
11	584	31.3	368	1	CCR3_HUMAN	P49682 homo sapien
12	581.5	31.1	367	1	CCR3_MOUSE	O88410 mus musculu
13	570.5	30.6	360	1	CCR4_HUMAN	P51679 homo sapien
14	562	30.1	359	1	IL8B_RAT	P35407 rattus norv
15	551.5	29.5	355	1	IL8A_RABIT	P21109 oryctolagus
16	550.5	29.5	378	1	CKR7_MOUSE	P47774 mus musculu
17	548.5	29.4	356	1	IL8B_CANFA	O97571 canis famil
18	544.5	29.2	358	1	IL8B_RABIT	P35344 oryctolagus
19	544.5	29.2	360	1	IL8B_HUMAN	P25025 homo sapien
20	542.5	29.1	353	1	IL8B_PANTR	Q28807 pan troglod
21	541	29.0	353	1	IL8B_MACMU	Q28519 macaca mula
22	541	29.0	353	1	IL8B_GORGO	Q28422 gorilla gor
23	537.5	28.8	378	1	CKR7_HUMAN	P32248 homo sapien
24	536	28.7	369	1	CKR9_MOUSE	Q9wut7 mus musculu
25	535	28.7	374	1	CKR6_HUMAN	P51684 homo sapien
26	534	28.6	357	1	CKR9_HUMAN	P51686 homo sapien
27	533	28.5	355	1	C3X1_HUMAN	P49238 homo sapien
28	532	28.5	367	1	CKR6_MOUSE	O54689 mus musculu
29	530	28.4	360	1	CKD6_MOUSE	P51680 mus musculu
30	528	28.3	378	1	IL8A_GORGO	O08707 mus musculu
31	525.5	28.1	350	1	IL8A_PANTR	P55919 gorilla gor
32	524.5	28.1	350	1	IL8B_MOUSE	P55920 pan troglod
33	523.5	28.0	359	1		P35343 mus musculu

```
                              ALIGNMENTS
```

					ID	Accession	Species
34	522.5	28.0	372	1	CCR5_HUMAN	P32302	homo sapien
35	520.5	27.9	350	1	IL8A_HUMAN	P25024	homo sapien
36	517.5	27.7	374	1	CCR5_MOUSE	Q04683	mus musculu
37	512.5	27.5	360	1	IL8B_BOVIN	Q28003	bos taurus
38	511.5	27.4	353	1	CKR8_MOUSE	P56484	mus musculu
39	510	27.3	342	1	CCR6_HUMAN	O00574	homo sapien
40	510	27.3	342	1	CCR6_PANTR	Q9tv16	pan troglod
41	508.5	27.2	374	1	CCR5_RAT	P34997	rattus norv
42	508	27.2	382	1	CKD6_RAT	O09027	rattus norv
43	505.5	27.1	342	1	CCR6_CERAE	O18983	cercopithec
44	504	27.0	342	1	CCR6_MACNE	O19024	macaca neme
45	501	26.8	354	1	C3X1_RAT	P35411	rattus norv

```
RESULT 1
CCR4_MOUSE
ID   CCR4_MOUSE      STANDARD;      PRT;    359 AA.
AC   P70658; O09059; O09062; P70233; P70346;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   C-x-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (Stromal cell-
DE   derived factor 1 receptor) (SDF-1 receptor) (Fusin) (Leukocyte-derived
DE   seven transmembrane domain receptor) (LESTR) (Pre-B-cell-derived
DE   chemokine receptor) (PB-CKR).
GN   CXCR4 OR LESTR OR CMKAR4 OR SDF1R.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM CXCR4-B).
RC   STRAIN=C57BL/6J, and 129/Sv; TISSUE=Peritoneal exudate;
RX   MEDLINE=97113334; PubMed=8955194;
RA   Heesen M., Berman M.A., Benson J.D., Gerard C., Dorf M.E.;
RT   "Cloning of the mouse fusin gene, homologue to a human HIV-1
RT   co-factor.";
RL   J. Immunol. 157:5455-5460(1996).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM CXCR4-B).
RC   TISSUE=Pre-B cell;
RX   MEDLINE=97121456; PubMed=8962122;
RA   Nagasawa T., Nakajima K., Tachibana K., Iizasa H., Bleul C.C.,
RA   Yoshie O., Matsushima K., Yoshida N., Springer T.A., Kishimoto T.;
RT   "Molecular cloning and characterization of a murine pre-B-cell
RT   growth-stimulating factor/stromal cell-derived factor 1 receptor, a
RT   murine homolog of the human immunodeficiency virus 1 entry coreceptor
RT   fusin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:14726-14729(1996).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM CXCR4-B).
RC   STRAIN=129/Sv; TISSUE=Thymus;
RA   Schubel A., Burgstahler R., Lipp M.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS CXCR4-A AND CXCR4-B).
RC   STRAIN=C57BL/6J x CBA; TISSUE=Thymus;
RX   MEDLINE=97439495; PubMed=9235051;
RA   Moepps B., Frodl R., Rodewald H.-R., Baggiolini M., Gierschik P.;
RT   "Two murine homologues of the human chemokine receptor CXCR4 mediating
RT   stromal cell-derived factor 1alpha activation of Gi2 are
RT   differentially expressed in vivo.";
RL   Eur. J. Immunol. 27:2102-2112(1997).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORMS CXCR4-A AND CXCR4-B).
RX   MEDLINE=97255574; PubMed=9103415;
RA   Heesen M., Berman M.A., Hoepken U.E., Gerard N.P., Dorf M.E.;
RT   "Alternate splicing of mouse fusin/CXC chemokine receptor-4: stromal
RT   cell-derived factor-1alpha is a ligand for both CXC chemokine
RT   receptor-4 isoforms.";
```

```
RL   J. Immunol. 158:3561-3564 (1997).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM CXCR4-B).
RC   STRAIN=C57BL/6; TISSUE=Thymus;
RA   Suzuki G., Nakata Y., Uzawa A., Shirasawa T., Saito T., Mita K.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM CXCR4-B).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bonk S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=99095114; PubMed=9879064;
RA   Prodl R., Gierschik P., Moepps B.;
RT   "Genomic organization and expression of the CXCR4 gene in mouse and
RT   man: absence of a splice variant corresponding to mouse CXCR4-B in
RT   human tissues.";
RL   J. Recept. Signal Transduct. Res. 18:321-344(1998).
RN   [9]
RP   FUNCTION.
RX   MEDLINE=98295994; PubMed=9634237;
RA   Tachibana K., Hiroca S., Iizasa H., Yoshida H., Kawabata K.,
RA   Kataoka Y., Kitamura Y., Matsushima K., Yoshida N., Nishikawa S.-I.,
RA   Kishimoto T., Nagasawa T.;
RT   "The chemokine receptor CXCR4 is essential for vascularization of the
RT   gastrointestinal tract.";
RL   Nature 393:591-594(1998).
RN   [10]
RP   FUNCTION.
RX   MEDLINE=98295995; PubMed=9634238;
RA   Zou Y.-R., Kottmann A.H., Kuroda M., Taniuchi I., Littman D.R.;
RT   "Function of the chemokine receptor CXCR4 in haematopoiesis and in
RT   cerebellar development.";
RL   Nature 393:595-599(1998).
RN   [11]
RP   DEVELOPMENTAL STAGE.
RC   STRAIN=ICR;
RX   MEDLINE=99410349; PubMed=10479460;
RA   McGrath K.E., Koniski A.D., Maltby K.M., McGann J.K., Palis J.;
RT   "Embryonic expression and function of the chemokine SDF-1 and its
RT   receptor, CXCR4.";
RL   Dev. Biol. 213:442-456(1999).
CC   -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC       signal by increasing the intracellular calcium ions level.
CC       Involved in B-cell lymphopoiesis, bone-marrow myelopoiesis and in
CC       cardiac ventricular septum formation. Plays also an essential role
CC       in vascularization of the gastrointestinal tract, probably by
CC       regulating vascular branching and/or remodelling processes in
CC       endothelial cells. Involved in cerebellar neuronal layer
CC       formation, preventing premature migration of proliferating granule
CC       cells from the external granule layer inwards. In the CNS, could
CC       mediate hippocampal-neuron survival.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=CXCR4-B; Synonyms=LESTR-B;
CC         IsoId=P70658-1; Sequence=Displayed;
CC       Name=CXCR4-A; Synonyms=LESTR-A;
CC         IsoId=P70658-2; Sequence=VSP_001891;
CC   -!- TISSUE SPECIFICITY: Lymphocytes, macrophages, neutrophils,
CC       microglial cells and astrocytes. Found in spleen, thymus, bone
CC       marrow, lymph nodes and, at lower levels in brain, small
CC       intestine, stomach and kidney. CXCR4-A is predominant in all
CC       tissues tested.
CC   -!- DEVELOPMENTAL STAGE: High expression during embryonic development
CC       does not seem to be associated with the differentiation of any
CC       particular cell type, but is widely utilized when there is a
CC       requirement for cell movement. Frequently associated with less
CC       differentiated cell types and down-regulated with subsequent
CC       differentiation. Detected in sites with haemopoietic potential:
CC       the yolk sac (7.5, 8.5 and 12.5 dpc) and fetal liver (12.5 dpc).
CC       During gastrulation, at 7.2 to 7.8 dpc, expressed in the mesoderm
CC       and the definitive endoderm. As gastrulation pattern fades (8.5
CC       dpc), expression in the mesoderm is down-regulated, while it
CC       becomes predominant in neural ectoderm. Endodermal expression is
CC       retained in the foregut and later in a subset of foregut
CC       derivatives, including the stomach (10.5 dpc), the cystic ducts of
CC       the gallbladder and the lung epithelium (12.5 dpc). In neuronal
CC       tissue: at 10.5 and 12.5 dpc, expressed in the dorsal root
CC       ganglia, in the ventral mantle layer of the spinal cord (or basal
CC       plates), in the hindbrain. At 14.5 dpc, expression more tightly
CC       confined to the neural epithelium lining the ventricular space and
CC       to the external granular layer of the ventral rhombic lip (the
CC       developing cerebellum). Expressed in the outpocketing of the
CC       diencephalic floor at 10.5 dpc and in the developing thalamus and,
CC       to a lesser extent, the developing hypothalamus. At 14.5 dpc,
CC       restricted to the region where thalamus and hypothalamus meet.
CC       Detected in a discrete band of cells at the edge of the olfactory
CC       bulb. In the vascular system: expressed in the endothelium of
CC       numerous blood vessels, but not all, at 10.5, 11.5 and 12.5 dpc,
CC       such as vitelline/umbilical vessels, cardiac ventricular wall
CC       capillaries, facial vessels and, at 14.5 pdc, in the vasculature
CC       of the herniated gut. Expression seems to be associated with
CC       expanding vascular networks. In the heart development, expressed
CC       at 10.5 dpc in the precursor to the aortopulmonary (AP) septum. At
CC       12.5 dpc, detected in the AP septum at the base of the outflow
CC       tract and in the atrioventricular valves. Detected in craniofacial
CC       ectoderm from 10.5 to 14.5 dpc. At 10.5 and 11.5 dpc, expressed in
CC       the Rathke's pouch.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U59760; AAB07725.1; -.
DR   EMBL; U65580; AAC52953.1; -.
DR   EMBL; D87747; BAA13451.1; -.
DR   EMBL; Z80111; CAB02201.1; -.
DR   EMBL; Z80112; CAB02202.1; -.
DR   EMBL; X99581; CAA67893.1; -.
DR   EMBL; X99582; CAA67894.1; -.
DR   EMBL; AB000803; BAA19187.1; -.
DR   EMBL; BC031665; AAH31665.1; -.
DR   MGD; MGI:109563; Cxcr4.
DR   GO; GO:0007420; P:brain development; IMP.
DR   GO; GO:0030334; P:regulation of cell migration; IMP.
DR   GO; GO:0042098; P:T-cell proliferation; IMP.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
```

```
Query Match     100.0%;   Score 1867;  DB 1;  Length 359;
Best Local Similarity  100.0%;  Pred. No. 2.2e-115;
Matches 359;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGL  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGL  60

Qy   61 VILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTV 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTV 120

Qy  121 NLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVGVWIPALLLTIPDFIFADVS 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 NLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVGVWIPALLLTIPDFIFADVS 180

Qy  181 QGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQ 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 QGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQ 240

Qy  241 KRKALKTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFF 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 KRKALKTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFF 300

Qy  301 HCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 HCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
```

RESULT 2
CCR4_RAT
```
ID   CCR4_RAT      STANDARD;       PRT;       349 AA.
AC   O08565;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (Stromal cell-
DE   derived factor 1 receptor) (SDF-1 receptor) (Fusin) (Leukocyte-derived
DE   seven transmembrane domain receptor) (LESTR).
GN   CXCR4 OR CMKAR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Spleen;
RA   Harrison J.K., Salafranca M.N.;
RT   "Molecular cloning of rat CXCR4.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for the C-x-C chemokine SDF-1. Transduces a
CC       signal by increasing the intracellular calcium ions level.
CC       Involved in B-cell lymphopoiesis, bone-marrow myelopoiesis and in
CC       cardiac ventricular septum formation. Plays also an essential role
CC       in vascularization of the gastrointestinal tract, probably by
CC       regulating vascular branching and/or remodelling processes in
CC       endothelial cells. Involved in cerebellar neuronal layer
CC       formation, preventing premature migration of proliferating granule
CC       cells from the external granule layer inwards. In the CNS, could
CC       mediate hippocampal-neuron survival (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in neurons and in astrocytes.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
```

```
DR   EMBL; U90610; AAB50408.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT   DOMAIN         1     36   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      37     60   1 (POTENTIAL).
FT   DOMAIN        61     76   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      77     96   2 (POTENTIAL).
FT   DOMAIN        97    107   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     108    129   3 (POTENTIAL).
FT   DOMAIN       130    151   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     152    172   4 (POTENTIAL).
FT   DOMAIN       173    197   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     198    217   5 (POTENTIAL).
FT   DOMAIN       218    237   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     238    258   6 (POTENTIAL).
FT   DOMAIN       259    282   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     283    302   7 (POTENTIAL).
FT   DOMAIN       303    349   CYTOPLASMIC (POTENTIAL).
FT   DISULFID     106    183   BY SIMILARITY.
FT   MOD_RES        4      4   SULFATION (POTENTIAL).
FT   MOD_RES       18     18   SULFATION (POTENTIAL).
FT   CARBOHYD       8      8   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   349 AA;  39334 MW;  7E0789A605C60C09 CRC64;
```

```
Query Match      93.4%;   Score 1744.5;  DB 1;  Length 349;
Best Local Similarity   95.5%;  Pred. No. 2.2e-107;
Matches 338;  Conservative  4;  Mismatches  7;  Indels  5;  Gaps  1;

Qy    6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVM  65
         :||||||||||||||||||||||||||| :|:||||||||||||||||||||||||||||
Db    1 MEIYTSDNYSEEVGSGDYDSNKEPCFRDENENFNRIFLPTIYPIIFLTGIVGNGLVILVM  60

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 120

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
         ||||||||||||||||||| :|| |||||:|||||||||||||||||||| ||||:
Db  121 VLILAFISLDRYLAIVHATNSQARKLLAEKAVYVGVWIPALLLTIPDIIFADV-----S 175

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  176 QGDGRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 235

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
         ||||||||||||||||||||||||||||| ||||||:||:||||||||||||||||||||
Db  236 KTTVILILAFFACWLPYYVGISIDSFILLEVIKQGCEFESVVHKWISITEALAFFHCCLN 295

Qy  306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  296 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 349
```

RESULT 3
CCR4_PAPAN
```
ID   CCR4_PAPAN    STANDARD;       PRT;       352 AA.
AC   P56491;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (SDF-1 receptor)
DE   (Stromal cell-derived factor 1 receptor) (Fusin).
GN   CXCR4.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Benton P.A., Timanus D.K., Shearer M.H., Lee D.R., Kennedy R.C.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC       signal by increasing the intracellular calcium ions level.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the  Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use by  non-profit institutions as  long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------
DR   EMBL; AF031089; AAC63831.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT   DOMAIN        1     39       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     40     63       1 (POTENTIAL).
FT   DOMAIN       64     79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     80     99       2 (POTENTIAL).
FT   DOMAIN      100    110       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    111    132       3 (POTENTIAL).
FT   DOMAIN      133    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       4 (POTENTIAL).
FT   DOMAIN      176    200       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    201    220       5 (POTENTIAL).
FT   DOMAIN      221    240       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    241    261       6 (POTENTIAL).
FT   DOMAIN      262    285       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    286    305       7 (POTENTIAL).
FT   DOMAIN      306    352       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES      21     21       SULFATION (POTENTIAL).
FT   CARBOHYD     11     11       N-LINKED (GLCNAC...) (POTENTIAL).
FT   DISULFID    109    186       BY SIMILARITY.
SQ   SEQUENCE   352 AA;  39751 MW;  468E542E1851265A CRC64;

Query Match             90.3%;  Score 1686.5; DB 1; Length 352;
Best Local Similarity   90.7%;  Pred. No. 1.3e-103;
Matches  321;   Conservative  16;  Mismatches  12;  Indels  5;  Gaps  1;

Qy      6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVM 65
            :||||||||::||:|||||||||||:|:::|||||||||||||||:||||||||||||||
Db      4 ISIYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy     66 GYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
            ||||||||||||||||||||||||||:||||||:|||:|| ||||||||||||||||||
Db     64 GYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy    126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVGVWIPALLLTIPDFIPADVSQGDIS 185
            |||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db    124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVVGVWIPALLLTIPDFIPASV----S 178

Qy    186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIISKLSHSKGHQKRKAL 245
            :|||||:|::|| :|||:|||||||||||||||||:||::|||||||||||||||||||
Db    179 EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIISKLSHSKGHQKRKAL 238

Qy    246 KTTVLILAFFACWLPYYVVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
            ||||||||||||||||||||||||||:||||||||:|| ||||||||||||||||||||||
Db    239 KTTVLILAFFACWLPYYIGISIDSFILLEIIKQGCEFRNTVHKWISITEALAFFHCCLN 298

Qy    306 PILYAFLGAKFKSSAQHALNSMRGSSLKILSKGRGGHSSVSTESESSSSFHSS 359
            ||||||||||||||||||||||||||||::||||||||||||||:|||||||||
Db    299 PILYAFLGAKFKTSAQHALTSVSRQSSLKILSKGRGGHSSVSTESESSSSFHSS 352
```

```
RESULT 4
CCR4_MACFA
ID   CCR4_MACFA    STANDARD;    PRT;   352 AA.
AC   Q28474;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (SDF-1 receptor)
DE   (Stromal cell-derived factor 1 receptor) (Fusin) (LESTR).
GN   CXCR4.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tatsumi M., Takahashi H.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for the C-x-C chemokine SDF-1. Transduces a
CC       signal by increasing the intracellular calcium ions level.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the  Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit institutions as  long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------
DR   EMBL; D86579; BAA13126.1; -.
DR   PIR; G00048; G00048.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT   DOMAIN        1     39       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     40     63       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       64     79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     80     99       2 (POTENTIAL).
FT   DOMAIN      100    110       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    111    132       3 (POTENTIAL).
FT   DOMAIN      133    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       4 (POTENTIAL).
FT   DOMAIN      176    200       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    201    220       5 (POTENTIAL).
FT   DOMAIN      221    240       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    241    261       6 (POTENTIAL).
FT   DOMAIN      262    285       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    286    305       7 (POTENTIAL).
FT   DOMAIN      306    352       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES      21     21       SULFATION (POTENTIAL).
FT   CARBOHYD     11     11       N-LINKED (GLCNAC...) (POTENTIAL).
FT   DISULFID    109    186       BY SIMILARITY.
SQ   SEQUENCE   352 AA;  39753 MW;  432DA6C11B59EF8A CRC64;

Query Match             90.0%;  Score 1679.5; DB 1; Length 352;
Best Local Similarity   90.1%;  Pred. No. 3.8e-103;
Matches  319;   Conservative  17;  Mismatches  13;  Indels  5;  Gaps  1;

Qy      6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVM 65
            :||||||||::||:|||||||||||:|:::||||||||||||||:||||||||||||||||
Db      4 ISIYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy     66 GYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
            ||||||||||||||||||||||||||:||||||:|||:||.||||||||||::||||||
Db     64 GYQKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123
```

Qy   126 VLIIAFISLDRYLAIVHATNSQRPRKLLAEKAVVYGVWIPALLLTIPDFIFADVSQGDIS 185
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYGVWIPALLLTIPDFIFASV-----S 178

Qy   186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLLPGIVILSCYCIIISKLSHSKGHQKRKAL 245
           ||||||| ||:|| || :||||||||||||||||||||||||||||||||||||||
Db   179 EAADRYICDRFYPNDLWVVVFQFQHIMVGLLPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246 KTTVILIAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
           |||:||||||||||||||||| |||| | :|:||||||||||||||||||:||||||
Db   239 KTTVILIAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALGFFHCCLN 298

Qy   306 PILYAFLGAKFKSSAQHALNSMSRGGSLKILSKGKRGGHSSVSTESESSSFHS 359
           ||||||||||||:|:|||||| :|:||||||||||||||||||||||||||
Db   299 PILYAFLGAKFKTSAQHALTSVSRGGSLKILSKGKRGGHSSVSTESESSSFHS 352

RESULT 5
CCR4_HUMAN
ID   CCR4_HUMAN   STANDARD;   PRT;   352 AA.
AC   P30991; O60835; P55438; Q9UKN2;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (Stromal cell-
DE   derived factor 1 receptor) (SDF-1 receptor) (Fusin) (Leukocyte-derived
DE   seven transmembrane domain receptor (LESTR) (LCR1) (FB22) (NPYRL)
DE   (HM89) (CD184 antigen).
GN   CXCR4.
OS   Homo sapiens (Human), and
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606, 9598;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human; TISSUE=Lung;
RX   MEDLINE=93319629; PubMed=8329116;
RA   Herzog H., Hort Y.J., Shine J., Selbie L.A.;
RT   "Molecular cloning, characterization, and localization of the human
RT   homolog to the reported bovine NPY Y3 receptor: lack of NPY binding
RT   and activation.";
RL   DNA Cell Biol. 12:465-471(1993).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human; TISSUE=Fetal brain;
RX   MEDLINE=94052833; PubMed=8234909;
RA   Jazin E.E., Yoo H., Blomqvist A.G., Yee F., Weng G., Walker M.W.,
RA   Salon J., Larhammar D., Wahlestedt C.R.;
RT   "A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its
RT   human homologue, confers neither NPY binding sites nor NPY
RT   responsiveness on transfected cells.";
RL   Regul. Pept. 47:247-258(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human; TISSUE=Fetal spleen;
RX   MEDLINE=93315164; PubMed=8325644;
RA   Federsppiel B., Melhado I.G., Duncan A.M., Delaney A.D.,
RA   Schappert K.T., Clark-Lewis I., Jirik F.R.;
RT   "Molecular cloning of the cDNA and chromosomal localization of the
RT   gene for a putative seven-transmembrane segment (7-TMS) receptor
RT   isolated from human spleen.";
RL   Genomics 16:707-712(1993).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human; TISSUE=Monocytes;
RX   MEDLINE=94103215; PubMed=8276799;
RA   Loetscher M., Geiser T., O'Reilly T., Zwahlen R., Baggiolini M.,
RA   Moser B.;
RT   "Cloning of a human seven-transmembrane domain receptor, LESTR, that
RT   is highly expressed in leukocytes.";
RL   J. Biol. Chem. 269:232-237(1994).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=94092629; PubMed=7505609;
RA   Nomura H., Nielsen B.W., Matsushima K.;
RT   "Molecular cloning of cDNAs encoding a LD78 receptor and putative
RT   leukocyte chemotactic peptide receptors.";
RL   Int. Immunol. 5:1239-1249(1993).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION OF ITS HIV-1
RP   CORECEPTOR FUNCTION.
RC   SPECIES=Human;
RX   MEDLINE=96217947; PubMed=8629022;
RA   Feng Y., Broder C.C., Kennedy P.E., Berger B.A.;
RT   "HIV-1 entry cofactor: functional cDNA cloning of a seven-
RT   transmembrane, G protein-coupled receptor.";
RL   Science 272:872-877(1996).
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human; TISSUE=Peripheral blood leukocytes;
RX   MEDLINE=98136183; PubMed=9468539;
RA   Wegner S.A., Ehrenberg P.K., Chang G., Dayhoff D.E., Sleeker A.L.,
RA   Michael N.L.;
RT   "Genomic organization and functional characterization of the chemokine
RT   receptor CXCR4, a major entry co-receptor for human immunodeficiency
RT   virus type 1.";
RL   J. Biol. Chem. 273:4754-4760(1998).
RN   [8]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human;
RX   MEDLINE=98258970; PubMed=9559023;
RA   Caruz A., Samsom M., Alonso J.M., Alcami J., Baleux F.,
RA   Virelizier J.L., Parmentier M., Arenzana-Seisdedos F.;
RT   "Genomic organization and promoter characterization of human CXCR4
RT   gene.";
RL   FEBS Lett. 426:271-278(1998).
RN   [9]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human;
RX   MEDLINE=99408510; PubMed=10480633;
RA   Xiao L., Weiss S.H., Qari S.H., Rudolph D., Zhao C., Denny T.N.,
RA   Hodge T., Lal R.B.;
RT   "Partial resistance to infection by R5X4 primary HIV type 1 isolates
RT   in an exposed-uninfected individual homozygous for CCR5 32-base pair
RT   deletion.";
RL   AIDS Res. Hum. Retroviruses 15:1201-1208(1999).
RN   [10]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human; TISSUE=Peripheral blood lymphocytes;
RX   MEDLINE=99095114; PubMed=9879064;
RA   Prodl R., Gierschik P., Moepps B.;
RT   "Genomic organization and expression of the CXCR4 gene in mouse and
RT   man: absence of a splice variant corresponding to mouse CXCR4-B in
RT   human tissues.";
RL   J. Recept. Signal Transduct. Res. 18:321-344(1998).
RN   [11]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   SPECIES=Human; TISSUE=Neutrophils;
RX   MEDLINE=99384048; PubMed=10452968;
RA   Gupta S.K., Pillarisetti K.;
RT   "CXCR4-Lo: molecular cloning and functional expression of a novel
RT   human CXCR4 splice variant.";
RL   J. Immunol. 163:2368-2372(1999).
RN   [12]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human; TISSUE=Lung;
RA   Warren C.N., Aronstam R.S., Sharma S.V.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [13]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   SPECIES=Human; TISSUE=Colon;
RX   MEDLINE=22388257; PubMed=12477932;

RA Hoxie J.A.;

Query Match   89.9%;   Score 1678.5;   DB 1;   Length 352;
Best Local Similarity   90.1%;   Pred. No. 4.5e-103;
Matches 319;   Conservative 19;   Mismatches 11;   Indels 5;   Gaps 1;

```
Qy      6  VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVM   65
           ||||||||||:|||||||||:||||||:|||||:|||||||||||||||||||||||||||
Db      4  ISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVM   63

Qy     66  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHIIYTVNLYSS  125
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db     64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy    126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVVGWIPALLLTIPDFIFADVSQGDIS  185
           |||||||||||||||||||||:||||||||||||||:||||||||||||||||||:---
Db    124  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVVGWIPALLLTIPDFIFANV----S  178

Qy    186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
           ||||||:||||:||||||||||||||||||||||||:||||||||||||||||||||||
Db    179  EADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILEIIKQGCEFENTVHKWISITEE  238

Qy    246  KTTVLILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
           |||||||||||||||||:||||||:||||||::|:
Db    239  KTTVLILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEEALAFFHCCLN  298

Qy    306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
           ||||||||||||:||||||:||||:||||||||||||||||||||||||||||
Db    299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

RESULT 6
ID   CCR4_MACMU   STANDARD;   PRT;   352 AA.
AC   P79394; O02745; O46428;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (SDF-1 receptor)
DE   (Stromal cell-derived factor 1 receptor) (Fusin) (LESTR).
GN   CXCR4.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Indian macaque;
RX   MEDLINE=97213934; PubMed=9060623;
RA   Chen Z., Zhou P., Ho D.D., Landau N.R., Marx P.A.;
RT   "Genetically divergent strains of simian immunodeficiency virus use
RT   CCR5 as a coreceptor for entry.";
RL   J. Virol. 71:2705-2714(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97268687; PubMed=9108095;
RA   Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
RA   Sharron M., Samson M., Lu Z.-H., Clements J.E., Murphey-Corb M.,
RA   Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
RT   "Differential utilization of CCR5 by macrophage and T cell tropic
RT   simian immunodeficiency virus strains.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Chinese;
RX   MEDLINE=98252393; PubMed=9591719;
RA   Pretet J.-L., Guillet J.-G., Butor C.;
RT   "New widespread CXCR4 allele in rhesus macaques does not predict
RT   subspecies or clinical evolution.";
RL   AIDS Res. Hum. Retroviruses 14:639-641(1998).
CC   -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC   signal by increasing the intracellular calcium ions level.

RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [14]
RP   SEQUENCE FROM N.A.
RC   SPECIES=P.troglodytes;
RX   MEDLINE=98090115; PubMed=9430250;
RA   Pretet J.-L., Zerbib A.C., Girard M., Guillet J.-G., Butor C.;
RT   "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1.";
RL   AIDS Res. Hum. Retroviruses 13:1583-1587(1997).
RN   [15]
RP   SULFATION.
RC   SPECIES=Human;
RA   Farzan M., Mirzabekov T., Kolchinsky P., Wyatt R., Cayabyab M.,
RA   Gerard N.P., Gerard C., Sodroski J., Choe H.;
RT   "Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1
RT   entry.";
RL   Cell 96:667-676(1999).
RN   [16]
RP   FUNCTION.
RX   MEDLINE=96351077; PubMed=8752280;
RA   Bleul C.C., Farzan M., Choe H., Parolin C., Clark-Lewis I.,
RA   Sodroski J., Springer T.A.;
RT   "The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and
RT   blocks HIV-1 entry.";
RL   Nature 382:829-833(1996).
RN   [17]
RP   FUNCTION.
RX   MEDLINE=96351078; PubMed=8752281;
RA   Oberlin E., Amara A., Bachelerie F., Bessia C., Virelizier J.-L.,
RA   Arenzana-Seisdedos F., Schwartz O., Heard J.-M., Clark-Lewis I.,
RA   Legler D.F., Loetscher M., Baggiolini M., Moser B.;
RT   "The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents
RT   infection by T-cell-line-adapted HIV-1.";
RL   Nature 382:833-835(1996).
RN   [18]
RP   ERRATUM.
RA   Oberlin E., Amara A., Bachelerie F., Bessia C., Virelizier J.-L.,
RA   Arenzana-Seisdedos F., Schwartz O., Heard J.-M., Clark-Lewis I.,
RA   Legler D.F., Loetscher M., Baggiolini M., Moser B.;
RL   Nature 384:288-288(1996).
RN   [19]
RP   CHARACTERIZATION OF ITS HIV-1 CORECEPTOR FUNCTION.
RX   MEDLINE=97002453; PubMed=8849450;
RA   Lapham C.K., Ouyang J., Chandrasekhar B., Nguyen N.Y., Dimitrov D.S.,
RA   Golding H.;
RT   "Evidence for cell-surface association between fusin and the CD4-gp120
RT   complex in human cell lines.";
RL   Science 274:602-605(1996).
RN   [20]
RP   CHARACTERIZATION OF ITS HIV-2 RECEPTOR FUNCTION.
RX   MEDLINE=97083584; PubMed=8929542;
RA   Endres M.J., Clapham P.R., Marsh M., Ahuja M., Turner J.D.,
RA   McKnight A., Thomas J.F., Stoebenau-Haggarty B., Choe S., Vance P.J.,
RA   Wells T.N.C., Power C.A., Sutterwala S.S., Doms R.W., Landau N.R.,

```
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U73740; AAC51159.1; -.
DR   EMBL; U93311; AAB54116.1; -.
DR   EMBL; AF001928; AAC39641.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT   DOMAIN        1     39       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     40     63       1 (POTENTIAL).
FT   DOMAIN       64     79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     80     99       2 (POTENTIAL).
FT   DOMAIN      100    110       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    111    132       3 (POTENTIAL).
FT   DOMAIN      133    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       4 (POTENTIAL).
FT   DOMAIN      176    200       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    201    220       5 (POTENTIAL).
FT   DOMAIN      221    240       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    241    261       6 (POTENTIAL).
FT   DOMAIN      262    285       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    286    305       7 (POTENTIAL).
FT   DOMAIN      306    352       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES      21     21       SULFATION (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC...) (POTENTIAL).
FT   DISULFID    109    186       BY SIMILARITY.
FT   CONFLICT     67     67       K -> E (IN REF. 2).
FT   CONFLICT    214    214       D -> V (IN REF. 2 AND 3).
FT   CONFLICT    348    348       S -> N (IN REF. 2).
SQ   SEQUENCE   352 AA;  39739 MW;  EC415E4820699C3F CRC64;

Query Match             89.8%;  Score 1676.5;  DB 1;  Length 352;
Best Local Similarity   90.1%;  Pred. No. 6e-103;
Matches  319;  Conservative  17;  Mismatches  13;  Indels  5;  Gaps  1;

Qy      6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHNRIFLPTIYFIIFLTGIVGNGLVILVM 65
          :|||||||:||||||||||||||||:||:|||:||||||||||||::||||||||||||
Db      4 ISIYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy     66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          ||||||||||||||||||||||||||||||||||:|:||:|||||||||||||||||||||
Db     64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHIIYTVNLYSS 123

Qy    126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||:|---S 178
Db    124 VLILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWIPALLLTIPDFIFASV----S 178

Qy    186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          ||||||||:|::|:|| ||:||||||||||||:||||||||||||||||||||||||||
Db    179 EADDRVICDRFYPNDLWVVVFQFQHIMVGLILPGIDILSCYCIIISKLSHSKGHQKRKAL 238

Qy    246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFRSIVHKWISITEALAFFHCCLN 305
          |||||||||||||||||||||||||||:|||::||||:|||:|||||||||||||||||
Db    239 KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy    306 PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGHSSVSTESESSSFHSS 359
          ||||||||||||::||||:||||||||||||||||||||||||||||||||||
Db    299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGHSSVSTESESSSFHSS 352
```

```
RESULT 7
CCR4_CERTO
ID   CCR4_CERTO   STANDARD;   PRT;  352 AA.
AC   O62747;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (SDF-1 receptor)
DE   (Stromal cell-derived factor 1 receptor) (Fusin) (LESTR).
GN   CXCR4.
OS   Cercocebus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercocebus.
OX   NCBI_TaxID=9531;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98321155; PubMed=9656999;
RA   Chen Z., Gettie A., Ho D.D., Marx P.A.;
RT   "Primary SIVsm isolates use the CCR5 coreceptor from sooty mangabeys
RT   naturally infected in west Africa: a comparison of coreceptor usage
RT   of primary SIVsm, HIV-2, and SIVmac.";
RL   Virology 246:113-124(1998).
CC   -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC       signal by increasing the intracellular calcium ions level.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF051906; AAC39834.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT   DOMAIN        1     39       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     40     63       1 (POTENTIAL).
FT   DOMAIN       64     79       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     80     99       2 (POTENTIAL).
FT   DOMAIN      100    110       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    111    132       3 (POTENTIAL).
FT   DOMAIN      133    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       4 (POTENTIAL).
FT   DOMAIN      176    200       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    201    220       5 (POTENTIAL).
FT   DOMAIN      221    240       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    241    261       6 (POTENTIAL).
FT   DOMAIN      262    285       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    286    305       7 (POTENTIAL).
FT   DOMAIN      306    352       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES      21     21       SULFATION (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC...) (POTENTIAL).
FT   DISULFID    109    186       BY SIMILARITY.
SQ   SEQUENCE   352 AA;  39648 MW;  4027432B8032F87E CRC64;

Query Match             89.4%;  Score 1668.5;  DB 1;  Length 352;
Best Local Similarity   89.5%;  Pred. No. 2e-102;
Matches  317;  Conservative  19;  Mismatches  13;  Indels  5;  Gaps  1;

Qy      6 VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHNRIFLPTIYFIIFLTGIVGNGLVLVM 65
          :|||||||:|||||||||||||||:||:||||||||||||||::||||||||||||||
Db      4 ISIYTSDNYTEEMGSGDYDSIKEPCFREKNAHFNRIFLPTIYSIIFLTGIVGNGLVLVM 63

Qy     66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
```

```
DR   EMBL; AJ009816; CAA08839.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
```

```
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT   DOMAIN        1     40       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     41     64       1 (POTENTIAL).
FT   DOMAIN       65     80       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     81    100       2 (POTENTIAL).
FT   DOMAIN      101    111       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    112    133       3 (POTENTIAL).
FT   DOMAIN      134    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    176       4 (POTENTIAL).
FT   DOMAIN      177    201       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    202    221       5 (POTENTIAL).
FT   DOMAIN      222    241       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    242    262       6 (POTENTIAL).
FT   DOMAIN      263    286       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    287    306       7 (POTENTIAL).
FT   DOMAIN      307    353       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES      22     22       SULFATION (POTENTIAL).
FT   CARBOHYD     11     11       N-LINKED (GLCNAC...) (POTENTIAL).
FT   DISULFID    110    187       BY SIMILARITY.
FT   CONFLICT     67     67       Q -> H (IN REF. 3).
FT   CONFLICT    263    263       D -> E (IN REF. 3).
SQ   SEQUENCE   353 AA;   39935 MW;  EA2BD46068A6C05B CRC64;

Query Match            88.0%;     Score 1643;  DB 1;  Length 353;
Best Local Similarity  89.0%;     Pred. No. 9.4e-101;
Matches  314;  Conservative  18;  Mismatches  15;  Indels  6;  Gaps  2;

Qy     8  IYTSDNYSE-EVGSGDYDSNKEPGFRDENVHFENRIFLPFTYIYFIIFLTGIVGNGLVILVMG 66
          ||||||:  ::||||||| ||||||||::|||:|||:||||||:||:||||||||||||||||
Db     6  IYPSDNYTEDDLGSGDYDSMKEPCFREENAHFENRIFLPTVYSIIFLTGIVGNGLVILVMG 65

Qy    67  YQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSV 126
          |||||||||||||||||||||||||||||||||| :|||||||||||||:||||||||||
Db    66  YQKKLRSMTDKYRLHLSVADLLFVLTLPFWAVDAVANWYFGKFLCKAVVIYTVNLYSSV 125

Qy   127  LILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQ 186
          |||||||||||||||||||||||||||||:||||||||||||||||||:||::
Db   126  LILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV----RE 180

Qy   187  GDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALK 246
          |||||||||||||| : |||||||||||||||||||||||||||||||||||||:|||||
Db   181  ADGRYICDRFYPSDSWLVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGYQKRKALK 240

Qy   247  TTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNP 306
          |||||||||||||||||||||||||||| : ||||||||||||:||||||||||||||||
Db   241  TTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFESTVHKWISITEALAFFHCCLNP 300

Qy   307  ILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          |||||||||||:: ||||||| :||||||:||||||||||||||||||||||||
Db   301  ILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 353
```

RESULT 9

```
CCR4_BOVIN
ID   CCR4_BOVIN    STANDARD;    PRT;   353 AA.
AC   P25930;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (SDF-1 receptor)
DE   (Stromal cell-derived factor 1 receptor) (Fusin) (Leukocyte-derived
DE   seven transmembrane domain receptor) (LESTR) (LCR1).
GN   CXCR4.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
```

```
Db    64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGKFLCKAVHIIYTVNLYSS 123

Qy   126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS 185
          |||||||||||||||||||||||:|||||||:||||||||||||||||:||:
Db   124  VLILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWIPALLLTIPGFIFASV----S 178

Qy   186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          : |||||:|||| :|:|: ||||||||||||||||||||||||||||||||||||||||||
Db   179  EADDRFICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy   246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
          |||||||||||||||||||||||||||| : ||||||:||:|||||||||||||||||||
Db   239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN 298

Qy   306  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||: :||||| :||||||:||||||||||||||||||||||||||
Db   299  PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352
```

RESULT 8

```
CCR4_FELCA
ID   CCR4_FELCA    STANDARD;    PRT;   353 AA.
AC   P56498; O02700; P79172;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (SDF-1 receptor)
DE   (Stromal cell-derived factor 1 receptor) (Fusin) (LESTR).
GN   CXCR4.
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX   NCBI_TaxID=9685;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97404646; PubMed=9261358;
RA   Willett B.J., Picard L., Hosie M.J., Turner J.D., Adema K.,
RA   Clapham P.R.;
RT   "Shared usage of the chemokine receptor CXCR4 by the feline and human
RT   immunodeficiency viruses.";
RL   J. Virol. 71:6407-6415(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Willett B.J.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lerner D.L., Elder J.H.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99399239; PubMed=10470253;
RA   Kovacs E.M., Baxter G.D., Robinson W.F.;
RT   "Feline peripheral blood mononuclear cells express message for both
RT   CXC and CC type chemokines.";
RL   Arch. Virol. 144:273-285(1999).
CC   -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC       signal by increasing the intracellular calcium ions level.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
```

```
DR   EMBL; U63558; AAC48852.1; -.
DR   EMBL; U92795; AAB51765.1; -.
```

```
Db    4 IRIFTSDNYTEDDLGSGDYDSMKEPCFREENAHFNRIFLPTTYSIIFLTGIVGNGLVILV  63
           : |:||||:|: ::||||||:: ||||||||||| |||:||||||||:||:|
Qy   65 MGYQKKLRSMTDKYRLRHLSVADLLFVTLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYS 124
        |||||||||||||||||||||||||| ||||||||:||:|:||||||||||:||||||||
Db   64 MGYQKKLRSMTDKYRLRHLSVADLLFVLTLPFWAVDAVANWYFGKFLCKAVHVIYTVNLYS 123
Qy  125 SVILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGWIPALLLTIPDFIFADVSQGDI    184
        |||||||||||||||||||||| |||||||:|||||| ||:||||||:||||     ::
Db  124 SVILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGWLPAVLLTIPDLIFA-----DI    178
Qy  185 SQGDDRYICDRLYPDSLWMVVFQFQHIMVGLLLPGIVILSCYCIIISKLSHSKGHQKRKA  244
         :|: ||||||:||.|||:|||||||||:||||||||||||||||||||||||| |||:|
Db  179 KEVDERYICDRFYPSDLWVVFQFQHIVVGLLLPGIVILSCYCIIISKLSHSKGYQKRRA   238
Qy  245 LKTTVILILAFFACWLPYVVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCL  304
        |||||||||:||||||||||:||||||||| :|:.|||:|||||||||||||||||||||
Db  239 LKTTVILILTFFACWLPYVIGISIDSFILLEIIQQGCEFESTVHKWISITEALAFFHCCL  298
Qy  305 NPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGHSSVSTESESSSFHSS       359
        ||||||||||||| ||||||.| ||||||||||||||||||||||||||||||||
Db  299 NPILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGHSSVSTESESSSFHSS       353
```

OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Locus coeruleus;
RX   MEDLINE=92100053; PubMed=1661837;
RA   Rimland J., Xin W., Sweetnam P., Saijoh K., Nestler E.J., Duman R.S.;
RT   "Sequence and expression of a neuropeptide Y receptor cDNA.";
RL   Mol. Pharmacol. 40:869-875(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Werling D.;
RT   "Role of chemokines in respiratory syncytial virus infection.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SHOWS THAT IT IS NOT A NPY3-R.
RX   MEDLINE=94052833; PubMed=8234909;
RA   Jazin E.E., Yoo H., Blomqvist G., Yee F., Weng G., Walker M.W.,
RA   Salon J., Larhammar D., Wahlestedt C.R.;
RT   "A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its
RT   human homologue, confers neither NPY binding sites nor NPY
RT   responsiveness on transfected cells.";
RL   Regul. Pept. 47:244-258(1993).
CC   -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC   signal by increasing the intracellular calcium ions level.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Brain, heart, kidney, lung and liver.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -!- CAUTION: Was originally (Ref.1) thought to be a receptor for
CC   neuropeptide Y type 3 (NPY3-R).

DR   EMBL; M86739; -; NOT_ANNOTATED_CDS.
DR   EMBL; AF399642; AAK94452.1; -.
DR   PIR; S28787; S28787.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; TransmemBrane; Glycoprotein; Sulfation.
FT   DOMAIN   1   40   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   41   64   1 (POTENTIAL).
FT   DOMAIN   65   80   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   81   100   2 (POTENTIAL).
FT   DOMAIN   101   111   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   112   133   3 (POTENTIAL).
FT   DOMAIN   134   155   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   156   176   4 (POTENTIAL).
FT   DOMAIN   177   201   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   202   221   5 (POTENTIAL).
FT   DOMAIN   222   241   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   242   262   6 (POTENTIAL).
FT   DOMAIN   263   286   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   287   306   7 (POTENTIAL).
FT   DOMAIN   307   353   CYTOPLASMIC (POTENTIAL).
FT   MOD_RES   22   22   SULFATION (POTENTIAL).
FT   CARBOHYD   11   11   N-LINKED (GLCNAC...) (POTENTIAL).
FT   DISULFID   110   187   BY SIMILARITY.
SQ   SEQUENCE   353 AA; 39938 MW; 42FFE5BC7545505E CRC64;

Query Match   87.4%; Score 1631; DB 1; Length 353;
Best Local Similarity   86.8%; Pred. No. 5.7e-100;
Matches 308; Conservative 26; Mismatches 15; Indels 6; Gaps 2;

Qy   6 VSIYTSDNYSE-EVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILV 64

RESULT 10
CCR4_SHEEP
ID   CCR4_SHEEP   STANDARD;   PRT;   192 AA.
AC   Q28553;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   C-X-C chemokine receptor type 4 (CXC-R4) (CXCR-4) (SDF-1 receptor)
DE   (Stromal cell-derived factor 1 receptor) (Fusin) (Leukocyte-derived
DE   seven transmembrane domain receptor) (LESTR) (Fragment).
GN   CXCR4 OR LESTR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypothalamus;
RA   Dyer C.J., Matteri R.L., Keisler D.H.;
RT   "Development of an ovine Y3 cDNA and expression of the Y3 receptor
RT   mRNA in the ovine hypothalamus and pituitary.";
RL   Abstr. - Soc. Neurosci. 21:1890-1890(1995).
CC   -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC   signal by increasing the intracellular calcium ions level.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- PTM: Sulfated (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -!- CAUTION: Was originally (Ref.1) thought to be a receptor for
CC   neuropeptide Y type 3 (NPY3-R).

DR   EMBL; U38942; AAA81347.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; TransmemBrane; Glycoprotein; Sulfation.
FT   NON_TER   <1
FT   DOMAIN   <1   29   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   30   53   1 (POTENTIAL).
```

RESULT 11

CCR3_HUMAN

```
ID   CCR3_HUMAN     STANDARD;      PRT;   368 AA.
AC   P49682; O15185; Q9P2T4; Q9P2T5;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   C-X-C chemokine receptor type 3 (CXC-R3) (CXCR-3) (CKR-L2) (CD183
DE   antigen).
GN   CXCR3 OR GPR9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=97188912; PubMed=9064356;
RA   Loetscher M., Gerber B., Loetscher P., Jones S.A., Piali L.,
RA   Clark-Lewis I., Baggiolini M., Moser B.;
RT   "Chemokine receptor specific for IP10 and mig: structure, function,
RT   and expression in activated T-lymphocytes.";
RL   J. Exp. Med. 184:963-969(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gutierrez J., Varona R., Zaballos A., Lind P., Marquez G.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Warren C.N., Aronstam R.S., Sharma S.V.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, Lung, and Testis;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Matsusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 5-368 FROM N.A.
RX   MEDLINE=96115583; PubMed=8666380;
RA   Marchese A., Heiber M., Nguyen T., Heng H.H.Q., Saldivia V.R.,
RA   Cheng R., Murphy P.M., Tsui L.-C., Shi X., Gregor P., George S.R.,
RA   O'Dowd B.F., Docherty J.M.;
RT   "Cloning and chromosomal mapping of three novel genes, GPR9, GPR10,
RT   and GPR14, encoding receptors related to interleukin 8, neuropeptide
RT   Y, and somatostatin receptors.";
RL   Genomics 29:335-344(1995).
RN   [6]
RP   SEQUENCE OF 278-368 FROM N.A., AND VARIANTS GLN-292 AND THR-363.
RX   MEDLINE=21040285; PubMed=11196695;
RA   Kato H., Tsuchiya N., Tokunaga K.;
RT   "Single nucleotide polymorphisms in the coding regions of human CXC-
RT   chemokine receptors CXCR1, CXCR2 and CXCR3.";
RL   Genes Immun. 1:330-337(2000).
RN   [7]
RP   LIGAND BINDING.
RC   TISSUE=Fetal astrocytes;
RX   MEDLINE=98290735; PubMed=9625760;
RA   Cole K.E., Strick C.A., Paradis T.J., Ogborne K.T., Loetscher M.,
RA   Gladue R.P., Lin W., Boyd J.G., Moser B., Wood D.E., Sahagan B.G.,
RA   Neote K.;
RT   "Interferon-inducible T cell alpha chemoattractant (I-TAC): A novel
RT   non-ELR CXC chemokine with potent activity on activated T cells
RT   through selective high affinity binding to CXCR3.";
RL   J. Exp. Med. 187:2009-2021(1998).
CC   -!- FUNCTION: Receptor for SCYB9/MIG, SCYB10/INP10 and SCYB11/ITAC.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -!- DATABASE: NAME=PROW; NOTE=PROW 1:78-83(2000);
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/guide/65053941_g.htm".
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X95876; CAA65126.1; -.
DR   EMBL; Z79783; CAB02143.1; ALT_INIT.
DR   EMBL; AY242128; AAO92295.1; -.
DR   EMBL; BC034403; AAH34403.1; -.
DR   EMBL; U32674; AAC50500.1; -.
DR   EMBL; AB032735; BAA92297.1; -.
DR   EMBL; AB032736; BAA92298.1; -.
DR   Genew; HGNC:4540; CXCR3.
DR   MIM; 600894; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004950; F:chemokine receptor activity; TAS.
DR   GO; GO:0006960; P:antimicrobial humoral response (sensu Inver. ..; TAS.
```

```
FT   DOMAIN       54     69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     70     89       2 (POTENTIAL).
FT   DOMAIN       90    100       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      101    122       3 (POTENTIAL).
FT   DOMAIN      123    144       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      145    165       4 (POTENTIAL).
FT   TRANSMEM    166    190       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    191   >192       5 (POTENTIAL).
FT   DISULFID     99    176       BY SIMILARITY.
FT   MOD_RES      11     11       SULFATION (POTENTIAL).
FT   NON_TER     192    192
FT   SEQUENCE    192 AA;  22178 MW;  A8BCFE303C52BD98 CRC64;
```

```
Query Match          47.4%;   Score 885.5;   DB 1;   Length 192;
Best Local Similarity  83.1%;   Pred. No. 1.7e-51;
Matches  167;   Conservative  14;   Mismatches  11;   Indels   9;   Gaps   2;

Qy     9  YTSDNYSEVGSSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGYQ  68
          ||     :|| ||||||:||||||||  :  ||||||| :||| |||||||||||||||||
Db     1  YTED----DLGSGDYDSMKEPCFREENAHFNRIFLPTVYSIIFLTGIVGNGLVILVMGYQ  56

Qy    69  KKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVLI 128
          |||||||||||||||||||||||||||||||:|| :|||:||||||:|||||||||||||
Db    57  KKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGQFLCKAVHVIYTVNLYSSVLI 116

Qy   129  LAFISLDRYLAIVHATNSQRPRKLLAEKAVVYGVWIPALLLTIPDFIFADVSQGDISQGD 188
          ||||||||||||||||||||||||||||  || ||:||:|||||| :|| |      |
Db   117  LAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWLPAVLLTIPDLLIFA-----DIKEAD 171

Qy   189  DRYICDRLYPDSLWMVVFQFQ 209
          :|||||:||| ||||:|||||
Db   172  ERYICDRFYPSDLWLVVFQFQ 192
```

DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0006935; P:chemotaxis; TAS.
DR   GO; GO:0007204; P:cytosolic calcium ion concentration elevation; TAS.
DR   InterPro; IPR004070; CXC_3 receptor.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01532; CXCCHMKINER3.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Antigen;
KW   Polymorphism.
FT   DOMAIN   1   53   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   54   80   1 (POTENTIAL).
FT   DOMAIN   81   89   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   90   110   2 (POTENTIAL).
FT   DOMAIN   111   125   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   126   147   3 (POTENTIAL).
FT   DOMAIN   148   169   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   170   189   4 (POTENTIAL).
FT   DOMAIN   190   212   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   213   233   5 (POTENTIAL).
FT   DOMAIN   234   255   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   256   277   6 (POTENTIAL).
FT   DOMAIN   278   298   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   299   321   7 (POTENTIAL).
FT   DOMAIN   322   368   CYTOPLASMIC (POTENTIAL).
FT   DISULFID   124   203   BY SIMILARITY.
FT   CARBOHYD   22   22   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   32   32   N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT   292   292   R -> Q.
FT   /FTId=VAR_016240.
FT   VARIANT   363   363   A -> T.
FT   /FTId=VAR_016241.
FT   CONFLICT   75   75   A -> R (IN REF. 2).
SQ   SEQUENCE   368 AA;   40659 MW;   F08A3B44B2BBAD04 CRC64;

Query Match   31.3%;   Score 584; DB 1; Length 368;
Best Local Similarity   35.2%;   Pred. No. 1.5e-31;
Matches 124; Conservative 67; Mismatches 139; Indels 22; Gaps   5;

Qy   12 DNYSEVGSGDYDSN---KEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVMGY 67
                :|:|||: | | :| :: | |||| :|:||:| |||||:|| |||||:| :|
Db   21 ENFSSSYDYGENESDSCCCTSPPCPQDFSLNFDRAFLPALYSLLFLLGLLGNGAVAAVLLS 80

Qy   68 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL 127
                |:: ::|: | :||| |||||||||||||||||||||| |||||:|| ||||:|||||
Db   81 RRTALSSTDTFLLHLAVADTLLVITLPLWAVDAAVQWVFGSGLCKVAGALFNINFYAGAL 140

Qy   128 ILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQG 187
                :|| |||||| : ||| ||||| |||||||||| ||||||||||| | | |
Db   141 LLACISFDRYLNIVHATQLYRRGPPARVTLTCLAVWGLCLLFALPDFIFLS-----AHH 194

Qy   188 DDRY---ICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKA 244
                |:| : |: :|: || |: :|| |||||||||||:|| || : : :|:| ::|
Db   195 DERLNATHCQYNFP-QVGRTALRVLQLVAGFLLPLLVMAYCYAHILAVLLVSRQRRLRA 253

Qy   245 LKTTVILIAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCL 304
                |||| |:|||| || :||| :| |:|: | :|:| ||| |||| || ||||||||
Db   254 MRLVVVVVAFALCWTPYHLVLVLVDILMDLGALARNCGRESRVDVAKSVTSGLGYMHCCL 313

Qy   305 NPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGRGGHSSVSTESESSSF 356
                ||||||:|| | | |: |: |:|| : :| |:|:|
Db   314 NPLLYAFVGVKFRERMWMLI------LRLGCPNRGLQRQPSSSRRDSSW 357

ID   CCR3_MOUSE   STANDARD;   PRT;   367 AA.
AC   O88410;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)

DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   C-X-C chemokine receptor type 3 (CXC-R3) (CXCR-3).
GN   CXCR3 OR CMKAR3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98318636; PubMed=9653165;
RA   Soto H., Wang W., Strieter R.M., Copeland N.G., Gilbert D.J.,
RA   Jenkins N.A., Hedrick J., Zlotnik A.;
RT   "The CC chemokine 6Ckine binds the CXC chemokine receptor CXCR3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:8205-8210(1998).
CC   -!- FUNCTION: Receptor for SCYB9/MIG, SCYB10/INP10 and SCYB11/ITAC (By
CC   similarity). Binds to SCYA21.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF045146; AAC40163.1; -.
DR   MGD; MGI:1272207; Cxcr3.
DR   InterPro; IPR004070; CXC_3_receptor.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR01532; CXCCHMKINER3.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein (POTENTIAL).
FT   DOMAIN   1   52   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   53   79   1 (POTENTIAL).
FT   DOMAIN   80   88   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   89   109   2 (POTENTIAL).
FT   DOMAIN   110   124   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   125   146   3 (POTENTIAL).
FT   DOMAIN   147   168   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   169   188   4 (POTENTIAL).
FT   DOMAIN   189   211   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   212   232   5 (POTENTIAL).
FT   DOMAIN   233   254   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   255   276   6 (POTENTIAL).
FT   DOMAIN   277   297   EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   298   320   7 (POTENTIAL).
FT   DOMAIN   321   367   CYTOPLASMIC (POTENTIAL).
FT   DISULFID   123   202   BY SIMILARITY.
FT   CARBOHYD   22   22   N-LINKED (GLCNAC. . .)   (POTENTIAL).
FT   CARBOHYD   32   32   N-LINKED (GLCNAC. . .)   (POTENTIAL).
SQ   SEQUENCE   367 AA;   41016 MW;   029FBB778E3CD4EA CRC64;

Query Match   31.1%;   Score 581.5; DB 1; Length 367;
Best Local Similarity   35.2%;   Pred. No. 2.2e-31;
Matches 122; Conservative 70; Mismatches 130; Indels 25; Gaps   7;

Qy   17 EVGSGDYD-SNKEPCFRDENVHFNRIFLPTIYPIIFLTGIVGNGLVILVMGYQKKLRSMT 75
                ||||| | || || || : : || | ::: :| |||||||||||||||| |::
Db   28 DYGENESDFSDSPPCPQDFSLNFDRTFLPALYSLLFLLGLLGNGAVAVLLSQRTALSST 87

Qy   76 DKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVLLIAFISLD 135
                |||:||:||||:|:|||| |||||| | :| |||||||: | ||| | :|| :||
Db   88 DTFLHLAVADVLVLTLPLWAVDAAVQWVFGPGLCKVAGALFNIFYAGAFLLACISFD 147

Qy   136 RYLAIVHATNSQR--PRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDDRYI- 192
                ||| ||||:| || |||||:| :   :|||||||||||:||||| |||: |
Db   148 RYLSIVHATQIYRRDPRVRVALTCIV--VWGLCLLFALPDFIY------LSANYDQRLN 198

```
Qy  193  ---CDRLYPDSLMVVFQPQHIMVGLILPGIVILSCYCIIISKLSHSKGHQXRKALXTTV 249
              :|   :| .:  : :: :|| .|:||. ||.:| :|  .:| ::|| .:
Db  199  ATHCQYNFP-QVGRTALRVLQLVAGFLLPLLVMAYCVAHILAVLLVSRGQRRFRAMRLVV 257

Qy  250  ILILAFFACWLPYYVGISIDSFILLGVIKQGCDFES:VHKWISITBALAFFHCCLNPILY 309
              |:|| |. |||:| ||||  | ||.  :|  | ::     :. |  |||||||||:|
Db  258  VVVAAFAVCWTPYHLVVLVDILMDVGVLARNCGRKSHVDVAKSVTSGMGYMHCCLNPLLY 317

Qy  310  AFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSF 356
              |::| ||: :          :|      |:       .:
Db  318  AFVGVKFRBKMWMLFTRIGRSD-------QRGPQRQPSSSRRESSW 356
```

RESULT 13

CKR4_HUMAN

```
ID   CKR4_HUMAN     STANDARD;      PRT;   360 AA.
AC   P51679; Q9ULY6; Q9ULY7;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   C-C chemokine receptor type 4 (C-C CKR-4) (CC-CKR-4) (CCR4)
DE   (K5-5).
GN   CCR4 OR CMKBR4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=95370289; PubMed=7642634;
RA   Power C.A., Meyer A., Nemeth K., Bacon K.B., Hoogewerf A.J.,
RA   Proudfoot A.E.I., Wells T.N.C.;
RT   "Molecular cloning and functional expression of a novel CC chemokine
RT   receptor cDNA from a human basophilic cell line.";
RL   J. Biol. Chem. 270:19495-19500(1995).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANTS VAL-130 AND SER-178.
RX   MEDLINE=21040311; PubMed=11196669;
RA   Kato H., Tsuchiya N., Izumi S., Miyamasu M., Nakajima T., Kawasaki H.,
RA   Hirai K., Tokunaga K.;
RT   "New variations of human CC-chemokine receptors CCR3 and CCR4.";
RL   Genes Immun. 1:97-104(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Kopatz S.A., Aronstam R.S., Sharma S.V.;
RT   "cDNA clones of human proteins involved in signal transduction
RT   sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   FUNCTION.
RX   MEDLINE=97313486; PubMed=9169480;
RA   Imai T., Baba M., Nishimura M., Kakizaki M., Takagi S., Yoshie O.;
RT   "The T cell-directed CC chemokine TARC is a highly specific
RT   biological ligand for CC chemokine receptor 4.";
RL   J. Biol. Chem. 272:15036-15042(1997).
RN   [5]
RP   FUNCTION.
RX   MEDLINE=98104168; PubMed=9430724;
RA   Imai T., Chantry D., Raport C.J., Wood C.L., Nishimura M., Godiska R.,
RA   Yoshie O., Gray P.W.;
RT   "Macrophage-derived chemokine is a functional ligand for the CC
RT   chemokine receptor 4.";
RL   J. Biol. Chem. 273:1764-1768(1998).
RN   [6]
RP   FUNCTION.
RX   MEDLINE=99394604; PubMed=10466728;
RA   Campbell J.J., Haraldsen G., Pan J., Rottman J., Qin S., Ponath P.,
RA   Andrew D.P., Warnke R., Ruffing N., Kassam N., Wu L., Butcher E.C.;
RT   "The chemokine receptor CCR4 in vascular recognition by cutaneous but
RT   not intestinal memory T cells.";
RL   Nature 400:776-780(1999).
RN   [7]
RP   FUNCTION, TISSUE SPECIFICITY, AND PHOSPHORYLATION.
RX   MEDLINE=2019238; PubMed=10754297;
RA   Inngjerdingen M., Damaj B., Maghazachi A.A.;
RT   "Human NK cells express CC chemokine receptors 4 and 8 and respond to
RT   thymus and activation-regulated chemokine, macrophage-derived
RT   chemokine, and I-309.";
RL   J. Immunol. 164:4048-4054(2000).
CC   -!- FUNCTION: High affinity receptor for the C-C type chemokines
CC       TARC/SCYA17 and MDC/SCYA22. The activity of this receptor is
CC       mediated by G(i) proteins which activate a phosphatidylinositol-
CC       calcium second messenger system. Can function as a chemoattractant
CC       homing receptor on circulating memory lymphocytes and as a
CC       coreceptor for some primary HIV-2 isolates. In the CNS, could
CC       mediate hippocampal-neuron survival.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in the thymus, in
CC       peripheral blood leukocytes, including T cells, mostly cd4+ cells,
CC       and basophils, and in platelets; at lower levels, in the spleen
CC       and in monocytes. Detected also in macrophages, il-2-activated
CC       natural killer cells and skin-homing memory T cells, mostly the
CC       ones expressing the cutaneous lymphocyte antigen (CLA). Expressed
CC       in brain microvascular and coronary artery endothelial cells.
CC   -!- PTM: IN NATURAL KILLER CELLS, SCYA22 BINDING INDUCES
CC       PHOSPHORYLATION ON YET UNDEFINED SER/THR RESIDUES, MOST PROBABLY
CC       BY BETA-ADRENERGIC RECEPTOR KINASES 1 AND 2.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics  and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long  as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X85740; CAA59743.1; -.
DR   EMBL; AB023888; BAA86965.1; -.
DR   EMBL; AB023889; BAA86966.1; -.
DR   EMBL; AB023890; BAA86967.1; -.
DR   EMBL; AB023891; BAA86968.1; -.
DR   EMBL; AB023892; BAA86969.1; -.
DR   EMBL; AY325539; AAP84352.1; -.
DR   PIR; A57160; A57160.
DR   Genew; HGNC:1605; CCR4.
DR   MIM; 604836; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004950; F:chemokine receptor activity; TAS.
DR   GO; GO:0006935; P:chemotaxis; TAS.
DR   GO; GO:0007204; P:cytosolic calcium ion concentration elevation; TAS.
DR   GO; GO:0006955; P:immune response; TAS.
DR   GO; GO:0006954; P:inflammatory response; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1-2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Polymorphism.
FT   DOMAIN        1     39       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     40     67       1 (POTENTIAL).
FT   DOMAIN       68     77       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     78     98       2 (POTENTIAL).
FT   DOMAIN       99    111       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    112    133       3 (POTENTIAL).
FT   DOMAIN      134    150       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      151    175       4 (POTENTIAL).
FT   TRANSMEM    176    206       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    207    226       5 (POTENTIAL).
FT   DOMAIN      227    242       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      243    267       6 (POTENTIAL).
FT   TRANSMEM    268    284       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      285    308       7 (POTENTIAL).
FT   TRANSMEM    309    360       CYTOPLASMIC (POTENTIAL).
```

```
RT      interleukin-8 receptors.";
RL      J. Biol. Chem. 271:32770-32776(1996).
CC  -!- FUNCTION: Receptor to interleukin-8, which is a powerful
CC      neutrophils chemocactic factor. Binding of IL-8 to the receptor
CC      causes activation of neutrophils. This response is mediated via a
CC      G-protein that activate a phosphatidylinositol-calcium second
CC      messenger system. This receptor binds to IL-8 with a high affinity
CC      and to GRO/MGSA and NAP-2 also with a high affinity.
CC  -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC  -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
DR   EMBL; X77797; CA54824.1; -.
DR   EMBL; D63584; BAA09797.1; -.
DR   EMBL; U70998; AAC52961.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Chemotaxis.
FT   DOMAIN        1     47       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     48     74       1 (POTENTIAL).
FT   DOMAIN       75     83       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     84    104       2 (POTENTIAL).
FT   DOMAIN      105    119       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    120    141       3 (POTENTIAL).
FT   DOMAIN      142    162       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    163    182       4 (POTENTIAL).
FT   DOMAIN      183    207       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    208    230       5 (POTENTIAL).
FT   DOMAIN      231    250       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    251    272       6 (POTENTIAL).
FT   DOMAIN      273    296       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    297    314       7 (POTENTIAL).
FT   DOMAIN      315    359       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     23     23       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    201    201       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    202    202       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    118    195       BY SIMILARITY.
SQ   SEQUENCE   359 AA;  40532 MW;  EF547326A074ABDD CRC64;

Query Match             30.1%;  Score 562;  DB 1;  Length 359;
Best Local Similarity   35.0%;  Pred. No. 4e.30;
Matches  130;  Conservative  67;  Mismatches  126;  Indels  48;  Gaps  10;

Qy    12 DNYS-EEVGSGYD---------SNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGN 58
          |:|:||| |||||:||||||||
Db     7 DNFSLEDFFSGDIDSYNYSSDDPFTLSDAAFC-PSANLDINRYAVVVIYVLVTLLSLVGN 65

Qy    59 GLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIY 118
          |||||||||| |:|| ||||
Db    66 SLVMLVILYNRSTCSVTDVYLLNLAIADLLFALTLPVWAASKVNGWIFGSFLCKVFSPLQ 125

Qy   119 TVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFAD 178
          |||||||||||::|:|| |:|:||
Db   126 EITFYSVLLLACISMDRYLAIVHATSTLIQKRHLV-KPVCITMWFLSLVLSLPIFIL-- 182

Qy   179 VSQGDISQGDDRYICDRLYPD----SLWMVVFQFQHIMVGLLILPGIVILSCYCIIISKL 233
          ||| |:|::||| |:|:|||||
Db   183 --RTTVKANPSTVVC--YENIGNNTSKWRVVLRILPQTYGFLLLIMLFCYGFTLRTL 237

Qy   234 SHSKGHQCRKALKTTVILLAFFACWLPYYVGISIDSFILLGVIGKQGCDFESIVHKWISI 293
```

```
FT   CARBOHYD    183    183       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    194    194       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    110    187       BY SIMILARITY.
FT   VARIANT     130    130       L -> V.
                                  /FTId=VAR_010669.
FT   VARIANT     178    178       C -> S.
                                  /FTId=VAR_010670.
SQ   SEQUENCE   360 AA;  41402 MW;  51EBE12AD1FAFABF CRC64;

Query Match             30.6%;  Score 570.5;  DB 1;  Length 360;
Best Local Similarity   37.4%;  Pred. No. 1.1e-30;
Matches  120;  Conservative  63;  Mismatches  125;  Indels  13;  Gaps  6;

Qy     1 MEPISVSIYTSDNYSEEVGSGD--YDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGN 58
          | :: || | |:|| |:|:|:||
Db     1 MNPTDIADTTLD--ESIYSNYYLYESIPKPCTKEGIKAFGELFLPPLYSLVFVFGLLGN 57

Qy    59 GLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIY 118
          ||||||:::|||||
Db    58 SVVVVLFKYKRLRSMTDVYLLNLAISDLLFVFSLPFWGYYAADQWVFGLGLCKMISKWMY 117

Qy   119 TVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFAD 178
          :||||||||||:|: :|::||
Db   118 LVGFYSGIFFVMLMSIDRYLAIVHAVFSLRARTLTYGVITSLATWSVAVFASLPGFLPST 177

Qy   179 VSQGDISQGDDRYICDRLYP--DSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHS 236
          : || | ::: |: :::||
Db   178 C---YTERNHTY-CKTKYSLNSTTWKVLSSLEINILGLVIPLGIMLFCYSMIIRTLQHC 232

Qy   237 KGHQCRKALKTTVILLAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEA 296
          |:|::: | ::: |:
Db   233 KNEKKNKAVKMIFAVVVLFLGFWTPYNIVLFLETIVELEVL--QDCTFERYLDYAIQATET 291

Qy   297 LAFFHCCLNPILYAFLGAKFK 317
          |||||||||||| |||
Db   292 LAFVHCCLNPIIYFFLGEKFR 312

RESULT 14
ID   IL8B_RAT       STANDARD;      PRT;   359 AA.
AC   P35407;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   High affinity interleukin-8 receptor B (IL-8R B) (CXCR-2) (GRO/MGSA
DE   receptor).
GN   IL8RB OR CXCR2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Lung;
RX   MEDLINE=97361587; PubMed=9218548;
RA   Gobl A.E., Huang M.R., Wang S., Zhou Y., Oeberg K.;
RT   "Molecular cloning and characterization of a cDNA encoding the rat
RT   interleukin-8 receptor.";
RL   Biochim. Biophys. Acta 1326:171-177(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=liver;
RA   Konishi K., Shibata F., Watanabe K., Tsurufuji S., Nakagawa H.,
RA   Fujioka M.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Lung, and Spleen;
RX   MEDLINE=97115810; PubMed=8955112;
RA   Dunstan C.-A.N., Salafranca M.N., Adhikari S., Xia Y., Feng L.,
RA   Harrison J.K.;
RT   "Identification of two rat genes orthologous to the human
```

```
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; TransmemBrane; Glycoprotein;
KW   Chemotaxis.
FT   DOMAIN         1     40       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      41     67       1 (POTENTIAL).
FT   DOMAIN        68     73       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      74     92       2 (POTENTIAL).
FT   DOMAIN        93    114       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     115    138       3 (POTENTIAL).
FT   DOMAIN       139    159       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     160    184       4 (POTENTIAL).
FT   DOMAIN       185    204       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     205    232       5 (POTENTIAL).
FT   DOMAIN       233    247       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     248    270       6 (POTENTIAL).
FT   DOMAIN       271    290       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     291    313       7 (POTENTIAL).
FT   DOMAIN       314    355       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD       7      7       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD      21     21       N-LINKED (GLCNAC...) (POTENTIAL).
FT   DISULFID     115    192       BY SIMILARITY.
FT   CONFLICT      90    111       DLLFALTMPIWAVSKEKGWIFG ->
FT                                 PAFCPDHAYLGRLQGKRLDFR (IN REF. 2).
FT   CONFLICT     146    147       HA -> QS (IN REF. 2).
FT   CONFLICT     204    204       R -> C (IN REF. 2).
FT   CONFLICT     287    288       DI -> EL (IN REF. 2).
SQ   SEQUENCE    355 AA;   40622 MW;   EFF49ACB9D1E0F21 CRC64;

Query Match
Best Local Similarity   29.5%;   Score 551.5;  DB 1;  Length 355;
                        35.0%;   Pred. No. 1.9e-29;
Matches  121;   Conservative  72;   Mismatches 142;   Indels  11;   Gaps  6;

Qy    12  DNYSEVGSGDYDSNKEPCFRDENVHPNRIFLPLTIYFIIFLTGIVGNGLVILVMGYQKKL   71
              :  :  ||  ::||| |:  |:.:|||.|||:||::|||::
Db    17  DEFANATGMPPVEKDYSPCLVVTQT-LNKVVVVIYALVFLSLLGNSLVMLVILYSRSN    75

Qy    72  RSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVLILAF  131
          |:|| ||:|:|||||| ::|:|:|||  :||  |||::||  :::||:||||
Db    76  RSVTDVYLLNLAMADLLFALTMPIWAVSKEKGWIFGTPLCKVVSLVKEVNFYSGILLLAC  135

Qy   132  ISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQGDDRY  191
          |:|||||||||||  || :: :|.||.||||:|| : | ||: | |
Db   136  ISVDRYLAIVHATRTLTQKRHLV-KPICLGIWALSLILSLPFFLFRQV---FSPNNSSP  190

Qy   192  IC--DRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKTTV  249
             :: || | ||  :||  |  ||: :| | ||::|:.|
Db   191  VCYEDLGHNTAKWRMVLRILPHTFGPILPLLVMLFCYGFTLRTLFQAHMGQKHRAMRVIF  250

Qy   250  ILLLAFPACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISTEALAFPHCCLNPILY   309
           :||| || |:|:|| :|| ||| :| :| || |:| |
Db   251  AVVLIFLCWLPYNLVLLADTLMRTHVIQETCQRRNDIDRALDATEILGFLHSCLNPIIY   310

Qy   310  AFLGAKFKSSAQHALNSMSRG-SSLKILSKGKRGGHSSVSTESESS  354
          || |: |:  ::|| |:|:|: ..:||.| :.|
Db   311  AFIGQNFRNGFLKML--AARGLISKEPLTHRVRTSYTSSSTNVPSN  354

Search completed: May 17, 2004, 22:16:44
Job time : 28 secs
```

```
Db   238  FKAHMGQKHRAMRVIFAVVLVFLLCWLPYNIVLFTDTLMRTKLIKETCERQNEINKALEA  297
              | :  ::: |  |   |:||||:|:|| | |:|:||.||:|:|
Qy   294  TEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSK-------GKRGGH  344
                 ||:|||:||  ::   ||   ||
Db   298  TEILGFLHSCLNPIIYAFIGQKFR---HGL------LKIMANYGLVSKEFLAKEGRP  345

Qy   345  SSVSTESESS  355
          :: : ::|
Db   346  SFVGSSSANTS  356

RESULT 15
IL8A_RABIT
ID   IL8A_RABIT     STANDARD;      PRT;    355 AA.
AC   P21109;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   High affinity interleukin-8 receptor A (IL-8R A) (CXCR-1).
GN   IL8RA OR CXCR1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91378994; PubMed=1898400;
RA   Beckmann M.P., Munger W.E., Kozlosky C., Vanden Bos T., Price V.,
RA   Lyman S., Gerard N.P., Gerard C., Cerretti D.P.;
RT   "Molecular characterization of the interleukin-8 receptor.";
RL   Biochem. Biophys. Res. Commun. 179:784-789(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Neutrophils;
RX   MEDLINE=91056034; PubMed=1700779;
RA   Thomas K.M., Pyun H.Y., Navarro J.;
RT   "Molecular cloning of the fMet-Leu-Phe receptor from neutrophils.";
RL   J. Biol. Chem. 265:20061-20064(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Neutrophils;
RX   MEDLINE=92148149; PubMed=1737938;
RA   Lee J., Kuang W.-J., Rice G.C., Wood W.I.;
RT   "Characterization of complementary DNA clones encoding the rabbit
RT   IL-8 receptor.";
RL   J. Immunol. 148:1261-1264(1992).
CC   -!- FUNCTION: Receptor to interleukin-8, which is a powerful
CC       neutrophils chemotactic factor. Binding of IL-8 to the receptor
CC       causes activation of neutrophils. This response is mediated via a
CC       G-protein that activate a phosphatidylinositol-calcium second
CC       messenger system.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Neutrophils.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -!- CAUTION: Was originally (Ref.2) thought to be the receptor for
CC       fMet-Leu-Phe (N-formyl peptide receptor).
```

```
DR   EMBL; M74240; AAA31375.1; -.
DR   EMBL; M58021; AAA31377.1; -.
DR   EMBL; M82873; AAA31376.1; -.
DR   PIR; A23669; A23669.
DR   PIR; JQ1231; JQ1231.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 17, 2004, 21:20:57 ; Search time 128 Seconds
          (without alignments)
          884.930 Million cell updates/sec

Title: US-09-367-052-2
Perfect score: 1867
Sequence: 1 MEPISVSIYTSDNYSEEVGS.........KRGGHSSVSTEBSSSSFHSS 359

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_25:*
                 1: sp_arChea:*
                 2: sp_bacteria:*
                 3: sp_fungi:*
                 4: sp_human:*
                 5: sp_invertebrate:*
                 6: sp_mammal:*
                 7: sp_mhc:*
                 8: sp_organelle:*
                 9: sp_phage:*
                10: sp_plant:*
                11: sp_rodent:*
                12: sp_virus:*
                13: sp_vertebrate:*
                14: sp_unclassified:*
                15: sp_rvirus:*
                16: sp_bacteriap:*
                17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1758.5	94.2	349	11	Q8VD47	Q8vd47 rattus norv
2	1692.5	90.7	352	6	Q7YS92	Q7ys92 tupaia bela
3	1691.5	90.6	352	6	Q8HZU1	Q8hzu1 callithrix
4	1689.5	90.5	352	6	Q8HZU0	Q8hzu0 saimiri sci
5	1679.5	90.0	352	6	O77488	O77488 cercopithec
6	1678.5	89.9	347	6	Q9MZM9	Q9mzm9 ateles pani
7	1677.5	89.9	347	6	Q9MZP7	Q9mzp7 presbytis s
8	1676.5	89.8	347	6	Q9MZP8	Q9mzp8 colobus pol
9	1676.5	89.8	347	6	Q9MZQ0	Q9mzq0 pygathrix n
10	1676.5	89.8	347	6	Q9MZM5	Q9mzm5 pithecia pi
11	1676.5	89.8	347	6	Q9MZP0	Q9mzp0 macaca assa
12	1676.5	89.8	347	6	Q9MZQ2	Q9mzq2 pygathrix r
13	1676.5	89.8	347	6	Q9MZP4	Q9mzp4 presbytis f
14	1676.5	89.8	347	6	Q9MZP9	Q9mzp9 nasalis lar
15	1676.5	89.8	347	6	Q9MZP5	Q9mzp5 presbytis p
16	1676.5	89.8	347	6	Q9MZQ3	Q9mzq3 pygathrix a
17	1676.5	89.8	347	6	Q9MZN8	macaca thib
18	1676.5	89.8	347	6	Q9MZN9	macaca arct
19	1676.5	89.8	347	6	Q9MZN7	macaca neme
20	1676.5	89.8	347	6	Q9MZM7	callimico g
21	1676.5	89.8	347	6	Q9MZQ1	pygathrix b
22	1676.5	89.8	352	6	Q9TSQ8	cercopithec
23	1674.5	89.7	347	6	Q9MZM1	eulemur mac
24	1674.5	89.7	347	6	Q9MZN0	alouatta se
25	1672.5	89.6	347	6	Q9MZP2	mandrillus
26	1672.5	89.6	352	6	Q9BDS5	macaca fasc
27	1671.5	89.5	347	6	Q9MZN4	hylobates 1
28	1670.5	89.5	347	6	Q9MZN6	hylobates c
29	1670.5	89.5	347	6	Q9MZN5	hylobates h
30	1670.5	89.5	347	6	Q9MZN1	gorilla gor
31	1670.5	89.5	347	6	Q9MZN2	pongo pygma
32	1670.5	89.5	352	4	Q9BXA0	homo sapien
33	1668.5	89.4	347	6	Q9MZM6	callicebus
34	1668.5	89.4	347	6	Q9MZP3	presbytis e
35	1668.5	89.4	347	6	Q9MZP6	presbytis j
36	1665.5	89.2	347	6	Q9MZN3	hylobates s
37	1663.5	89.1	347	6	Q9MZM8	callithrix
38	1661.5	89.0	347	6	Q9MZP1	mandrillus
39	1644.5	88.1	347	6	Q9MZM0	perodicticu
40	1638	87.7	346	6	Q9MZM2	nycticebus
41	1638	87.7	346	6	Q9MZM3	nycticebus
42	1638	87.7	346	6	Q9MZM4	nycticebus
43	1613	86.4	332	11	Q62973	rattus norv
44	1611.5	86.3	334	6	Q9N1P2	saguinus oe
45	1606.5	86.0	334	6	Q9N130	saimiri bol

ALIGNMENTS

RESULT 1
Q8VD47
ID   Q8VD47   PRELIMINARY;   PRT;   349 AA.
AC   Q8VD47;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4.
GN   CXCR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holtzman; TISSUE=Brain;
RA   Simen A.A., Miller R.J.;
RT   "Chemokine regulation of neuronal signaling and gp120 neurotoxicity.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF452185; AAL47855.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. ..; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE 349 AA; 39429 MW; 09D19860D3D2CB8A CRC64;

Query Match   94.2%; Score 1758.5; DB 11; Length 349;
Best Local Similarity   96.0%; Pred. No. 7.8e-150;
Matches 340; Conservative 4; Mismatches 5; Indels 5; Gaps 1;

```
Qy    6  VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM  65
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MEIYTSDNYSEEVGSGDYDSNKEPCFRDENENHFNRIFLPTIYFIFLTGIVGNGLVILVM  60

Qy   66  GYQKKLRSMTDKYRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   61  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  120

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
            ||||||||||||||||||||||||||||||||||||||||||||||||  :|
Db  121  VLIILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDIIFADV-----S  175

Qy  186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  176  QGDGRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  235

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
            |||||||||||||||||||||||||||||||| ::|||||||||||||||||||||||
Db  236  KTTVILILARFFACWLPYYVGISIDSFILLEVIKQGCEFESVVHKWISITEALARFFHCCLN  295

Qy  306  PILYAFLGAKFKSSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
            ||||||||||||||||:||||||||||||||||||||||||||||||||||
Db  296  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  349
```

RESULT 2
Q7YS92
ID   Q7YS92;
AC   Q7YS92;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chemokine receptor CXCR4.
OS   Tupaia belangeri chinensis.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.
OX   NCBI_TaxID=103897;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang M., Ben K.;
RT   "Cloning and sequence analysis of HIV-1 related genes of CD4, CXCR4
RT   and CCR5 from tree shrew.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY177628; AAO47588.2; -.
KW   Receptor.
SQ   SEQUENCE   352 AA; 39719 MW; 606B07A6D46A7D53 CRC64;

```
Query Match             90.7%;  Score 1692.5;  DB 6;  Length 352;
Best Local Similarity   91.2%;  Pred. No. 6.7e-144;
Matches  323;  Conservative  15;  Mismatches  11;  Indels  5;  Gaps  1;

Qy    6  VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM  65
            :||||||| :|| :|||||||||||||||:| ||||||||| :|||||||||||||
Db    4  ISIYTSDNYSEELGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVM  63

Qy   66  GYQKKLRSMTDKYRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
            ||||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||
Db   64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS  123

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
            |||||||||||||||||||||||||||||||||||||||||||||||| :|
Db  124  VLIILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S  178

Qy  186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
            :||:||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  179  EAADRYICDRFYPNDLWMVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
            ||||||||||||||||||||||||||||||||||||||| :||||||||||||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
            ||||||||||||||||:||||||||||||||||||||||||||||||||||
Db  299  PILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

RESULT 3
Q8HZU1
ID   Q8HZU1;
AC   Q8HZU1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chemokine receptor CXCR4.
OS   Callithrix jacchus (Common marmoset).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
OX   NCBI_TaxID=9483;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22174698; PubMed=12186836;
RA   LaBonte J.A., Babcock G.J., Patel T., Sodroski J.;
RT   "Blockade of HIV-1 Infection of New World Monkey Cells Occurs
RT   Primarily at the Stage of Virus Entry.";
RL   J. Exp. Med. 196:431-445(2002).
DR   EMBL; AF452612; AAN14528.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS02262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Receptor.
SQ   SEQUENCE   352 AA; 39802 MW; 8018A4500FC887C0 CRC64;

```
Query Match             90.6%;  Score 1691.5;  DB 6;  Length 352;
Best Local Similarity   90.7%;  Pred. No. 8.2e-144;
Matches  321;  Conservative  18;  Mismatches  10;  Indels  5;  Gaps  1;

Qy    6  VSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVM  65
            :|||||||||||||| :|||||||||:||||||||||||| :|||||||||||||||
Db    4  ISIYTSDNYTEEIGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVM  63

Qy   66  GYQKKLRSMTDKYRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS  125
            ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   64  GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGKFLCKAVHVIYTVNLYSS  123

Qy  126  VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDIS  185
            |||||||||||||||||||||||||||||||||||||||||||||||| :|
Db  124  VLIILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIPDFIFANV-----S  178

Qy  186  QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  245
            :||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  179  EADDRYICDRFYPNDLWMVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL  238

Qy  246  KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN  305
            |||||||||||||||||||||||||||||||||||||||::|||||||||||||||||||
Db  239  KTTVILILAFFACWLPYYIGISIDSFILLEIIRQGCEFENTVHKWISITEALAFFHCCLN  298

Qy  306  PILYAFLGAKFKSSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
            ||||||||||||||||:||||||||||||||||||||||||||||||||||
Db  299  PILYAFLGAKFKSSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  352
```

RESULT 4
Q8HZU0
ID   Q8HZU0;
AC   Q8HZU0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chemokine receptor CXCR4.
OS   Saimiri sciureus (Common squirrel monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX   NCBI_TaxID=9521;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22174698; PubMed=12186836;
RA   LaBonte J.A., Babcock G.J., Patel T., Sodroski J.;
RT   "Blockade of HIV-1 Infection of New World Monkey Cells Occurs
RT   Primarily at the Stage of Virus Entry.";
RL   J. Exp. Med. 196:431-445(2002).
DR   EMBL; AF452613; AAN14529.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Receptor.
SQ   SEQUENCE   352 AA;  39820 MW;  BF1A21FB0C8D4487 CRC64;

Query Match             90.5%;  Score 1689.5;  DB 6;  Length 352;
Best Local Similarity   90.7%;  Pred. No. 1.2e-143;
Matches  321;  Conservative  18;  Mismatches  10;  Indels  5;  Gaps  1;

Qy    6 VSIYTSDNYSEBVGSGGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
          |||||||||| |||| || |||||||||||||||||||||||| |||| ||||||||||||
Db    4 ISIYTSDNYTEEMGSGGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          ||||||||||||||||||||||||||||||||||||: : :|||||||||:|||||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGKFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYYVGWIPALLLTIPDFIFADVSQGDIS 185
          ||||||||||||||||||||||||||||||: ||||||||||||||||||||:-----S 178
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYYVGWIPALLLTIPDFIFANV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          :||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
          ||||||||||||||||||||||||||||| :  ::||||||||||||||||||||||||
Db  239 KTTVILILAFFACWLPYYIGISIDSFILLEIIRQGCBFENTVHKWISITEALAFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQHALNSMSRGGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||:|| :: : :||||||||||||||||||||||||||||||||||
Db  299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352


RESULT 5
O77488
ID   O77488           PRELIMINARY;      PRT;   352 AA.
AC   O77488;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CXCR4 receptor.
OS   Cercopithecus aethiops (Green monkey) (Grivet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=9534;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Murayama Y., Matsunaga S., Inoue-Murayama M.;
```

```
RT   "cDNA sequence of African green monkey CXCR-4 chemokine receptor
RT   gene.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB015943; BAA31327.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   352 AA;  39773 MW;  DDC9182E14540EB9 CRC64;

Query Match             90.0%;  Score 1679.5;  DB 6;  Length 352;
Best Local Similarity   90.4%;  Pred. No. 9.9e-143;
Matches  320;  Conservative  16;  Mismatches  13;  Indels  5;  Gaps  1;

Qy    6 VSIYTSDNYSEBVGSGGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
          |||||||||| |||| || |||||||||||||||||||||||| |||| ||||||||||||
Db    4 ISIYTSDNYTEEMGSGGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVM 63

Qy   66 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
          |||||||||||||||||||||||||||||||||||| |||: :|||||||||:|||||||
Db   64 GYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSS 123

Qy  126 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYYVGWIPALLLTIPDFIFADVSQGDIS 185
          |||||||||||||||||||||||||||||||||||||||||||||:||||||:-----S 178
Db  124 VLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYYVGWIPALLLTIPHFIFASV-----S 178

Qy  186 QGDDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 245
          :||||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||
Db  179 EADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL 238

Qy  246 KTTVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLN 305
          ||||||||||||||||||||||||||||||||  ::||||||||||||||||||||||||
Db  239 KTTVILILAFFACWLPYYIGISIDSFILLEIIRQGCBFENTVHKWISITEALAFFHCCLN 298

Qy  306 PILYAFLGAKFKSSAQHALNSMSRGGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          ||||||||||||:|| :: :||||||||||||||||||||||||||||||||||||
Db  299 PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 352


RESULT 6
Q9MZM9
ID   Q9MZM9           PRELIMINARY;      PRT;   347 AA.
AC   Q9MZM9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Ateles paniscus (Black spider monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Atelinae; Ateles.
OX   NCBI_TaxID=9510;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172234; AAF89354.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
```

DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER   1   1
SQ   SEQUENCE   347 AA; 39317 MW; B84037A04490A1D1 CRC64;

Query Match   89.9%; Score 1678.5; DB 6; Length 347;
Best Local Similarity   90.6%; Pred. No. 1.2e-142;
Matches 319; Conservative 18; Mismatches 10; Indels   5; Gaps   1;

Qy   8 IYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVMGY 67
          ||||||||:||||||||:|||||||:|:|||||||||||||||:|||||||||||||||
Db   1 IYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVMGY 60

Qy   68 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL 127
          |||||||||||||||||||||||||||||||:||::|||||||||||||:||||||||||
Db   61 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL 120

Qy   128 ILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQG 187
          |||||||||||||||||||||:||||||||::||||||||||||||||||:|||| :||
Db   121 ILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWIPALLLTIPDFIF-----SEA 175

Qy   188 DDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT 247
          |||||||||:|||||||||||||||||||||||||||||||:||||||||||||||||||
Db   176 DDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT 235

Qy   248 TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI 307
          ||||||||||||||||||||||:|||||||||:|||||||||||:|||||||||||||||
Db   236 TVILIAFFACWLPYYIGISIDSFILLEIIRQGCEFENTVHKWISITEALAFFHCCLNPI 295

Qy   308 LYAFLGAKFKSSAQHALNSMSRGGSLKILSKGKRGGHSSVSTESESSSSFHSS 359
          ||||||||:|||||||||:||||||||||||||||||||||||||||||:||
Db   296 LYAFLGAKFKTSAQHALTSVSRGGSLKILSKGKRGGHSSVSTESESSSSFHSS 347

RESULT 7
Q9MZP7
ID   Q9MZP7   PRELIMINARY;   PRT;   347 AA.
AC   Q9MZP7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Presbytis senex.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
OC   Presbytis.
OX   NCBI_TaxID=66055;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172216; AAF89336.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.

DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER   1   1
SQ   SEQUENCE   347 AA; 39191 MW; 2F69F0CA9CBC3878 CRC64;

Query Match   89.9%; Score 1677.5; DB 6; Length 347;
Best Local Similarity   90.3%; Pred. No. 1.5e-142;
Matches 318; Conservative 17; Mismatches 12; Indels   5; Gaps   1;

Qy   8 IYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVMGY 67
          ||||||||:||||||||:|||||||:|:|||||||||||||||:|||||||||||||||
Db   1 IYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVMGY 60

Qy   68 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL 127
          |||||||||||||||||||||||||||||||:||::|||||||||||||:||||||||||
Db   61 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL 120

Qy   128 ILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQG 187
          |||||||||||||||||||||:||||||||::|||||||||||||||||:|||: ::||
Db   121 ILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWIPALLLTIPDFIF-----GSVSEA 175

Qy   188 DDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT 247
          |||||||||:|||||||||||||||||||||||||||||||:||||||||||||||||||
Db   176 DDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT 235

Qy   248 TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI 307
          ||||||||||||||||||||||:|||||||||:|||||||||||:|||||||||||||||
Db   236 TVILIAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLNPI 295

Qy   308 LYAFLGAKFKSSAQHALNSMSRGGSLKILSKGKRGGHSSVSTESESSSSFHSS 359
          ||||||||:|||||||||:||||||||||||||||||||||||||||||:||
Db   296 LYAFLGAKFKTSAQHALTSVSRGGSLKILSKGKRGGHSSVSTESESSSSFHSS 347

RESULT 8
Q9MZP8
ID   Q9MZP8   PRELIMINARY;   PRT;   347 AA.
AC   Q9MZP8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Colobus polykomos.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
OC   Colobus.
OX   NCBI_TaxID=9572;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172215; AAF89335.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER   1   1
SQ   SEQUENCE   347 AA; 39205 MW; E96B6613654I515C CRC64;

Query Match   89.8%;   Score 1676.5;   DB 6;   Length 347;
Best Local Similarity   90.6%;   Pred. No. 1.8e-142;
Matches 319;   Conservative 16; Mismatches 12; Indels   5; Gaps   1;

```
Qy    8  IYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVNGLVILVMGY  67
         |||||| |||| ||| |||| ||||| | |||||||||||| ||||||||||||||||||
Db    1  IYTSDNYTEEMGSGGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVNGLVILVMGY  60

Qy   68  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL 127
         |||||||||||||||||||||||||||||||||: ::|||:|||||||:|||||||||||
Db   61  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL 120

Qy  128  ILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQG 187
         ||||||||||||||||||||:|||||||| ||||||||||||||||||||:    :
Db  121  ILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWIPALLLTIPDFIFASV-----SEA 175

Qy  188  DDRYICDRLYPDSLMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT 247
         |||||||| || ||:|||||||||||||||||||||||||||||||||||||||||||
Db  176  DDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT 235

Qy  248  TVILILAFFACWLPYYVGISIDSPILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI 307
         ||||||||||||||||:|||||| |||| ::|||:|| .||||||||||||||||||||
Db  236  TVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLNPI 295

Qy  308  LYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSSFHSS 359
         |||||||||| ||||||.:::||||.||||||||||||||||||||:|||||
Db  296  LYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 347
```

RESULT 9
Q9MZQ0

```
ID   Q9MZQ0          PRELIMINARY;     PRT;    347 AA.
AC   Q9MZQ0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Pygathrix nemaeus (Dove langur).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
OC   Pygathrix.
OX   NCBI_TaxID=54133;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172213; AAF89333.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER       1      1
SQ   SEQUENCE   347 AA;  39205 MW;  E96B66136541515C CRC64;
```

Query Match   89.8%;   Score 1676.5;   DB 6;   Length 347;
Best Local Similarity   90.6%;   Pred. No. 1.8e-142;
Matches 319;   Conservative 16; Mismatches 16; Indels 12; Gaps   1;

```
Qy    8  IYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVNGLVILVMGY  67
```

```
Db    1  IYTSDNYTEEMGSGGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVNGLVILVMGY  60

Qy   68  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL 127
         |||||||||||||||||||||||||||||||||: ::|||:|||||||:|||||||||||
Db   61  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL 120

Qy  128  ILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQG 187
         ||||||||||||||||||||:|||||||| ||||||||||||||||||||:    :
Db  121  ILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWIPALLLTIPDFIFASV-----SEA 175

Qy  188  DDRYICDRLYPDSLMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT 247
         |||||||| || ||:|||||||||||||||||||||||||||||||||||||||||||
Db  176  DDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT 235

Qy  248  TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI 307
         ||||||||||||||||:|||||| |||| ::|||:|| .||||||||||||||||||||
Db  236  TVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLNPI 295

Qy  308  LYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTEBESSSSFHSS 359
         |||||||||| ||||||.:::||||:||||||||||||||||||:||||||
Db  296  LYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 347
```

RESULT 10
Q9MZM5

```
ID   Q9MZM5          PRELIMINARY;     PRT;    347 AA.
AC   Q9MZM5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Pithecia pithecia (White-faced saki).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Pitheciinae;
OC   Pithecia.
OX   NCBI_TaxID=43777;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172238; AAF89358.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER       1      1
SQ   SEQUENCE   347 AA;  39290 MW;  E59B2070BF0F58E4 CRC64;
```

Query Match   89.8%;   Score 1676.5;   DB 6;   Length 347;
Best Local Similarity   90.6%;   Pred. No. 1.8e-142;
Matches 319;   Conservative 17; Mismatches 17; Indels   5; Gaps   1;

```
Qy    8  IYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVNGLVILVMGY  67
         |||||| |||| ||| |||| ||||| | |||||||||||| ||||||||||||||||||
Db    1  IYTSDNYTEEMGSGGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVNGLVILVMGY  60

Qy   68  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL 127
         |||||||||||||||||||||||||||||||||: ::|||:|||||||:|||||||||||
Db   61  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL 120
```

```
Qy  128  ILAFISLDRYLAIVHATNSQRPKLLAEKAVVGVWIPALLLTIPDFIFPADVSQGDISQG  187
                                                      :.:
Db  121  ILAFISLDRYLAIVHATTSQRPKLLAEKVVVGVWIPALLLTIPDFIFANV-----SEA  175

Qy  188  DDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT  247
         |||||:|:||||||||||||||||||||||||:|||||||||||:|:||:|||||||||||
Db  176  DDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT  235

Qy  248  TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI  307
         ||||||||||||||||||||||||||:||:::|.:||::||||||||||||||||||||||
Db  236  TVILILAFFACWLPYYIGISIDSFILLEIIRQGCEFENTVHKWISITEALAFFHCCLNPI  295

Qy  308  LYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  359
         |||.|||||:.|:||||.||::|||||||||||:|||||||||:||||||||
Db  296  LYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS  347

RESULT 11
Q9MZP0
ID   Q9MZP0;
AC   Q9MZP0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Macaca assamensis (Assam's macaque) (Assam's monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9551;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172223; AAF89343.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER   1   1
SQ   SEQUENCE   347 AA;   39205 MW;   E96B6613654151 5C CRC64;

Query Match   89.8%;   Score 1676.5;   DB 6;   Pred. No. 1.8e-142;   Length 347;
Best Local Similarity   90.6%;
Matches  319;   Conservative  16;   Mismatches  12;   Indels  5;   Gaps  1;

Qy    8  IYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVMGY  67
         |||||||:||.||||||||||:||:|||||.||:|||||:|::|||||||||||||||||
Db    1  IYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVMGY  60

Qy   68  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL  127
         |||||||||||||||||||||||||||||||:||:||||||||:||||||||||||||||
Db   61  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL  120

Qy  128  ILAFISLDRYLAIVHATNSQRPKLLAEKAVVGVWIPALLLTIPDFIFPADVSQGDISQG  187
                                                      :.:
Db  121  ILAFISLDRYLAIVHATTSQRPKLLAEKVVVGVWIPALLLTIPDFIFASV-----SEA  175

Qy  188  DDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT  247
         |||||:|:||||||||||||||||||||||||:|||||||||||:|:||:|||||||||||
Db  176  DDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT  235

Qy  248  TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI  307
         ||||||||||||||||||||||||||:||:::|.:||::||||||||||||||||||||||
Db  236  TVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLNPI  295

RESULT 12
Q9MZQ2
ID   Q9MZQ2;
AC   Q9MZQ2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Pygathrix roxellana (golden snub-nosed monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
OC   Pygathrix.
OX   NCBI_TaxID=61622;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172211; AAF89331.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER   1   1
SQ   SEQUENCE   347 AA;   39205 MW;   E96B6613654151 5C CRC64;

Query Match   89.8%;   Score 1676.5;   DB 6;   Pred. No. 1.8e-142;   Length 347;
Best Local Similarity   90.6%;
Matches  319;   Conservative  16;   Mismatches  12;   Indels  5;   Gaps  1;

Qy    8  IYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIFLTGIVGNGLVILVMGY  67
         |||||||:||.||||||||||:||:|||||.||:|||||:|::|||||||||||||||||
Db    1  IYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVMGY  60

Qy   68  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL  127
         |||||||||||||||||||||||||||||||:||:||||||||:||||||||||||||||
Db   61  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL  120

Qy  128  ILAFISLDRYLAIVHATNSQRPKLLAEKAVVGVWIPALLLTIPDFIFPADVSQGDISQG  187
                                                      :.:
Db  121  ILAFISLDRYLAIVHATNSQKPKLLAEKVVVGVWIPALLLTIPDFIFASV-----SEA  175

Qy  188  DDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT  247
         |||||:|:||||||||||||||||||||||||:|||||||||||:|:||:|||||||||||
Db  176  DDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT  235

Qy  248  TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI  307
         ||||||||||||||||||||||||||:||:::|.:||::||||||||||||||||||||||
Db  236  TVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISITEALAFFHCCLNPI  295
```

Qy   308 LYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          |||||||||| :|:|||||| ||||||||||||||||||||||||||||||||
Db   296 LYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 347

RESULT 13
Q9MZP4
ID   Q9MZP4   PRELIMINARY;   PRT;   347 AA.
AC   Q9MZP4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Presbytis francoisi.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
OC   Presbytis.
OX   NCBI_TaxID=33549;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172219; AAF89339.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER   1   1
SQ   SEQUENCE   347 AA; 39205 MW; E96B6613654151 5C CRC64;

Query Match
Best Local Similarity   89.8%; Score 1676.5; DB 6; Length 347;
Matches 319; Conservative 16; Mismatches 12; Indels 5; Gaps 1;

Qy   8 IYTSDNYSEEVGSGGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGY 67
           ||||||| ||| |||||||||:|||||| |||||||||:|::||||||||||||||||||
Db   1 IYTSDNYTEEMGSGGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVMGY 60

Qy   68 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL 127
          ||||||||||||||||||||||||||||||:|:|:||||||:||||||||:|||||||
Db   61 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL 120

Qy   128 ILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQG 187
          ||||||||||||||||||||| ::||||||::|||||||||||||||:||||||----SEA 175
Db   121 ILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWIPALLLTIPDFIFASV-----SEA 175

Qy   188 DDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVIKQGCDFESIVHKWISITEALAFFHCCLNPI 247
          ||||||||:||| || |||||||||||||||||:|||| |||:||||||||||||||||||
Db   176 DDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVLEIIKQGCEFNTVHKWISITEALAFFHCCLNPI 235

Qy   248 TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI 307
          ||||||||||||||||||||||||||: :|||||:||||||||||||||||||||||
Db   236 TVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFNTVHKWISITEALAFFHCCLNPI 295

Qy   308 LYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          |||||||||| :|:|||||| ||||||||||||||||||||||||||||||||
Db   296 LYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 347

RESULT 14
Q9MZP9
ID   Q9MZP9   PRELIMINARY;   PRT;   347 AA.
AC   Q9MZP9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Nasalis larvatus (Proboscis monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
OC   Nasalis.
OX   NCBI_TaxID=43780;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172214; AAF89334.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER   1   1
SQ   SEQUENCE   347 AA; 39205 MW; E96B6613654151 5C CRC64;

Query Match
Best Local Similarity   89.8%; Score 1676.5; DB 6; Length 347;
Matches 319; Conservative 16; Mismatches 12; Indels 5; Gaps 1;

Qy   8 IYTSDNYSEEVGSGGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVMGY 67
           ||||||| ||| |||||||||:|||||| |||||||||:|::||||||||||||||||||
Db   1 IYTSDNYTEEMGSGGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVMGY 60

Qy   68 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL 127
          ||||||||||||||||||||||||||||||:|:|:||||||:||||||||:|||||||
Db   61 QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL 120

Qy   128 ILAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGVWIPALLLTIPDFIFADVSQGDISQG 187
          ||||||||||||||||||||| ::||||||::|||||||||||||||:||||||----SEA 175
Db   121 ILAFISLDRYLAIVHATNSQKPRKLLAEKVVYVGVWIPALLLTIPDFIFASV-----SEA 175

Qy   188 DDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVIKQGCDFESIVHKWISITEALAFFHCCLNPI 247
          ||||||||:||| || |||||||||||||||||:|||| |||:||||||||||||||||||
Db   176 DDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVLEIIKQGCEFNTVHKWISITEALAFFHCCLNPI 235

Qy   248 TVILILAFFACWLPYYVGISIDSFILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI 307
          ||||||||||||||||||||||||||: :|||||:||||||||||||||||||||||
Db   236 TVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFNTVHKWISITEALAFFHCCLNPI 295

Qy   308 LYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 359
          |||||||||| :|:|||||| ||||||||||||||||||||||||||||||||
Db   296 LYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSSFHSS 347

RESULT 15
Q9MZP5
ID   Q9MZP5   PRELIMINARY;   PRT;   347 AA.
AC   Q9MZP5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Chemokine receptor CXCR4 (Fragment).
GN   CXCR4.
OS   Presbytis phayrei.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
OC   Presbytis.
OX   NCBI_TaxID=66054;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT   "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
RT   Primates.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF172218; AAF89338.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER       1     1
SQ   SEQUENCE   347 AA;  39205 MW;  E96B6136541515C CRC64;
```

```
Query Match              89.8%;  Score 1676.5;  DB 6;  Length 347;
Best Local Similarity    90.6%;  Pred. No. 1.8e-142;
Matches  319;  Conservative  16;  Mismatches  12;  Indels  5;  Gaps  1;

Qy     8  IYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLIGIVGNGLVILVMGY  67
          |||||| ::|::::|||||:||||||||:||||||||||||||||||||||||||||||
Db     1  IYTSDNYTEEMGSGDYDSIKEPCFREENAHFNRIFLPTIYSIIFLTGIVGNGLVILVMGY  60

Qy    68  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIIYTVNLYSSVL  127
          ||||||||||||||||||||||||||||||||||:|:||:||||||||||||||||||||
Db    61  QKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL  120

Qy   128  ILAFISLDRYLAIVHATNSQRPRKLLAEKAVVVGVWIPALLLTIPDFIFADVSQGDISQG  187
          ||||||||||||||||||||:|||||||:||||||||||||||||||||:
Db   121  ILAFISLDRYLAIVHATNSQKPRKLLAEKVVVVGVWIPALLLTIPDFIFASV-----SEA  175

Qy   188  DDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT  247
          ||||||||:|: :|:||||||||||||||||||||||:|||||||||||||||||||||
Db   176  DDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKALKT  235

Qy   248  TVILILAFFACWLPYYVGISIDSPILLGVIKQGCDFESIVHKWISITEALAFFHCCLNPI  307
          |||||||||||||||:|||||||||:|:||||||||||:|||||||||||||||||||||
Db   236  TVILILAFFACWLPYYIGISIDSPILLEIIKQGCEFNTVHKWISITEALAFFHCCLNPI  295

Qy   308  LYAFLGAKFKSSAQHALNSMSRGGSLKILSKGKRGGHSSVSTESESSSFHSS  359
          ||||||||||:|:||||:||||||||||||||||||||||||||||||||||
Db   296  LYAFLGAKFKTSAQHALTSVSRGGSLKILSKGKRGGHSSVSTESESSSFHSS  347
```

Search completed: May 17, 2004, 22:19:03
Job time : 131 secs

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 17, 2004, 04:56:59 ; Search time 2770.88 Seconds
                (without alignments)
                10714.982 Million cell updates/sec

Title: US-09-367-052-5
Perfect score: 685
Sequence: 1 ccatcctaatacgactcact.........gacatcagtcaggggggatga 685
Scoring table: IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 2167151 6995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
  1: gb_ba:*
  2: gb_htg:*
  3: gb_in:*
  4: gb_om:*
  5: gb_ov:*
  6: gb_pat:*
  7: gb_ph:*
  8: gb_pl:*
  9: gb_pr:*
  10: gb_ro:*
  11: gb_sts:*
  12: gb_sy:*
  13: gb_un:*
  14: gb_vi:*
  15: em_ba:*
  16: em_fun:*
  17: em_hum:*
  18: em_in:*
  19: em_mu:*
  20: em_om:*
  21: em_or:*
  22: em_ov:*
  23: em_pat:*
  24: em_ph:*
  25: em_pl:*
  26: em_ro:*
  27: em_sts:*
  28: em_un:*
  29: em_vi:*
  30: em_htg_hum:*
  31: em_htg_inv:*
  32: em_htg_other:*
  33: em_htg_mus:*
  34: em_htg_pln:*
  35: em_htg_rod:*
  36: em_htg_mam:*
  37: em_htg_vrt:*
  38: em_sy:*
  39: em_htgo_hum:*
  40: em_htgo_mus:*
  41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	685	100.0	1877	10	D87747	D87747 Mus musculu
2	636	92.8	1817	10	BC031665	BC031665 Mus muscu
3	631.2	92.1	1575	10	X99582	X99582 M.musculus
4	618.6	90.3	1809	10	AB000803	AB000803 Mouse mRN
5	595	86.9	1223	10	MMU59760	U59760 Mus musculu
6	548.4	80.1	1180	10	MMLCR12	Z80112 Mus musculu
7	545.8	79.7	1422	10	MMLCR13	Z80113 Mus musculu
8	545.8	79.7	3366	10	MMU65580	U65580 Mus musculu
9	541	79.0	3770	10	MMLESTRGN	X99581 M.musculus
10	541	79.0	175370	2	AC145861	AC145861 Pan trogl
11	541	79.0	185215	2	AC147556	AC147556 Mus muscu
12	495.8	72.4	528	10	MMLCR11	Z80111 Mus musculu
13	473.4	69.1	190149	2	AC122097	AC122097 Rattus no
14	472.4	69.0	1050	10	AF452185	AF452185 Rattus no
15	469.2	68.5	1050	10	RNU90610	U90610 Rattus norv
16	452.8	66.1	1126	9	AF019378	AF019378 Cercopith
17	446.8	65.2	1667	9	AK129916	AK129916 Homo sapi
18	446.8	65.2	1679	6	AX287116	AX287116 Sequence
19	446.8	65.2	1679	6	AX548790	AX548790 Sequence
20	446.8	65.2	1679	6	AX705078	AX705078 Sequence
21	445.8	65.1	1664	6	E37853	E37853 Anti-HIV ag
22	445.8	65.1	1664	6	AR225504	AR225504 Sequence
23	445.8	65.1	1664	6	AR380668	AR380668 Sequence
24	445.8	65.1	1664	6	HUMHM89	D10924 Homo sapien
25	445.8	65.1	1664	11	G28514	G28514 SHGC-31451
26	444.8	64.9	1637	9	HUMSTSR	M99293 Homo sapien
27	441.6	64.5	1645	9	HSNPYRLA	X71635 H.sapiens m
28	440.4	64.3	1059	9	AF031089	AF031089 Papio ham
29	439.2	64.1	1737	6	AR015970	AR015970 Sequence
30	439.2	64.1	1737	6	AR060748	AR060748 Sequence
31	439.2	64.1	1737	6	AR070433	AR070433 Sequence
32	439.2	64.1	1737	6	AR103430	AR103430 Sequence
33	439.2	64.1	1737	6	I13753	I13753 Sequence 4
34	439.2	64.1	1737	6	I24455	I24455 Sequence 3
35	438.8	64.1	1068	9	AB015943	AB015943 Chloroceb
36	437.6	63.9	1670	9	AX409496	AX409496 Sequence
37	437.6	63.9	1670	9	AX663641	AX663641 Sequence
38	437.6	63.9	1662	9	HUMGPCR	L06797 Human (clon
39	437.2	63.8	1662	9	BC020968	BC020968 Homo sapi
40	435.6	63.6	1078	9	AF291672	AF291672 Macaca fa
41	434	63.4	1059	9	AF452613	AF452613 Saimiri s
42	434	63.4	1059	9	MMU73740	U73740 Macaca mula
43	434	63.4	1078	9	D86579	D86579 Macaca fasc
44	433.8	63.3	1044	9	AF172212	AF172212 Pygathrix
45	433.8	63.3	1044	9	AF172238	AF172238 Pithecia

ALIGNMENTS

RESULT 1
D87747
LOCUS   D87747   1877 bp   mRNA   linear   ROD 07-FEB-1999
DEFINITION   Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION   D87747
VERSION   D87747.1 GI:1772445
KEYWORDS   murine CXCR-4.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (sites)
  AUTHORS   Nagasawa,T., Nakajima,T., Tachibana,K., Iizasa,H., Bleul,C.C.,
              Yoshie,O., Matsushima,K., Yoshida,N., Springer,T.A. and
              Kishimoto,T.

Pred. No. is the number of results predicted by chance to have a

```
Qy    361 ACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCA 420

Qy    421 TGGCTGACTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCACTCACACTGTCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGCTGACTGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCACACTACTGTCA 480

Qy    481 ACCTTCACAGCAGCGGTTCTCATCCTCGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCTTCACAGCAGCGGTTCTCATCCTCGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTG 540

Qy    541 TCCAGGCCACCAACAGTCAAGAGGCCAAGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCCAGGCCACCAACAGTCAAGAGGCCAAGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600

Qy    601 GCGTCTGGATCCAGCCCTCCTGACTATACCTGACTTCTTGCCGACGTCAGCC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCGTCTGGATCCAGCCCTCCTGACTATACCTGACTTCTTGCCGACGTCAGCC 660

Qy    661 AGGGGGACATCAGTCAGGGGATGA 685
          ||||||||||||||||||||||||
Db    661 AGGGGGACATCAGTCAGGGGATGA 685


RESULT 2
BC031665
LOCUS           BC031665         1817 bp    mRNA         linear   ROD 07-OCT-2003
DEFINITION      Mus musculus chemokine (C-X-C motif) receptor 4, mRNA (cDNA clone
                MGC:36266 IMAGE:3592479), complete cds.
ACCESSION       BC031665
VERSION         BC031665.1    GI:21618781
KEYWORDS        MGC.
SOURCE          Mus musculus (house mouse)
  ORGANISM      Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1  (bases 1 to 1817)
  AUTHORS       Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
                Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
                Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
                Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
                Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
                Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
                Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
                Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
                Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
                McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
                Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
                Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
                Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
                Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
                Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
                Dickson,M.C., Rodrigues,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
                Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
                Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE         Generation and initial analysis of more than 15,000 full-length
                human and mouse cDNA sequences
  JOURNAL       Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE       22388257
   PUBMED       12477932
REFERENCE       2  (bases 1 to 1817)
  AUTHORS       Strausberg,R.
  TITLE         Direct Submission
  JOURNAL       Submitted (06-JUN-2002) National Institutes of Health, Mammalian
                Gene Collection (MGC), Cancer Genomics Office, National Cancer
                Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
                USA
  REMARK        NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT         Contact: MGC help desk
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: Gilbert Smith, Ph.D.
                cDNA Library Preparation: Life Technologies, Inc.
```

The second column (left side of rotated page):

```
TITLE           Molecular cloning and characterization of a murine pre-B-cell
                growth-stimulating factor/stromal cell-derived factor 1 receptor, a
                murine homolog of the human immunodeficiency virus 1 entry
                coreceptor fusin
JOURNAL         Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)
MEDLINE         97121456
 PUBMED         8962122
REFERENCE       2  (bases 1 to 1877)
  AUTHORS       Nagasawa,T.T.
  TITLE         Molecular cloning of murine PBSF/SDF-1 receptor
  JOURNAL       Unpublished
REFERENCE       3  (bases 1 to 1877)
  AUTHORS       Nagasawa,T.
  TITLE         Direct Submission
  JOURNAL       Submitted (05-SEP-1996) Takashi Nagasawa, Research Institute, Osaka
                Medical Center, Department of Immunology; Murodocho 840, Izumi,
                Osaka 590-02, Japan (E-mail:immunol@osk.threewebnet.or.jp,
                Tel:0725-56-1220, Fax:0725-57-3021)
FEATURES             Location/Qualifiers
     source          1..1877
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="DW34"
                     /cell_type="pre-B cell"
                     /tissue_type="bone marrow"
     gene            1..1877
                     /gene="murine CXCR-4"
     CDS             120..1199
                     /gene="murine CXCR-4"
                     /function="PBSF/SDF-1 receptor"
                     /codon_start=1
                     /evidence=experimental
                     /product="murine CXCR-4"
                     /protein_id="BAA13451.1"
                     /db_xref="GI:1772446"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLLPFWAVDAMA
                     DWYPGKPLCKAVHIITVNLYSSVLLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                     GIVILSCYCIIISKLSHSKGHQKRALKTTVILLAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTSESSSSFHSS"
ORIGIN

Query Match         100.0%;     Score 685;  DB 10;  Length 1877;
Best Local Similarity   100.0%;  Pred. No. 5.8e-170;
Matches  685;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCCGGGCAGGTCAGGTAGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCCGGGCAGGTCAGGTAGCAG 60

Qy     61 TGACCCTCGAGCGGTTTGGTGCTCCGGTAACCACCACCACCACCGGCGTGTAGAGCGGAGTGTTGCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGACCCTCGAGCGGTTTGGTGCTCCGGTAACCACCACCACCGGCGTGTAGAGCGGAGTGTTGCCA 120

Qy    121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTTCTGAAGAAGTGGGGTCTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTTCTGAAGAAGTGGGGCTG 180

Qy    181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGA 240

Qy    241 TCTTCCTGCCACCATCACTTCATCATCATCTTCTTGACTCGGCATAGTCGGCAATGATTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCTTCCTGCCACCATCACTTCATCATCATCTTCTTGACTCGGCATAGTCGGCAATGATTGG 300

Qy    301 TGATCCTGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGATCCTGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 360
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
  Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 61 Row: d Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES   Location/Qualifiers
     source   1..1817
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="FVB/N"
                  /db_xref="taxon:10090"
                  /clone="MGC:36266 IMAGE:3592479"
                  /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                  model. 10 month old virgin mouse. Taken by biopsy."
                  /clone_lib="NCI_CGAP_Mam1"
                  /lab_host="DH10B"
                  /note="Vector: pCMV-SPORT6"
     gene   1..1817
                  /gene="Cxcr4"
                  /note="synonyms: LESTR, fusin, PB-CKR, PBSF/SDF-1, CD184"
                  /db_xref="LocusID:12767"
                  /db_xref="MGI:109563"
     CDS   78..1157
                  /codon_start=1
                  /product="chemokine (C-X-C motif) receptor 4"
                  /protein_id="AAH31665.1"
                  /db_xref="GI:21618782"
                  /db_xref="LocusID:12767"
                  /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                  TIYFIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                  DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                  GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLLLP
                  GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
                  IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                  ILSKGKRGHGHSSVSTESESSSFHSS"
     misc_feature 246..1004
                  /note="7tm_1; Region: 7 transmembrane receptor (rhodopsin
                  family)"
                  /db_xref="CDD:pfam00001"

ORIGIN

Query Match   92.8%; Score 636; DB 10; Length 1817;
Best Local Similarity 100.0%; Pred. No. 5.2e-157;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   50 GCAGGTAGCAGTGACCCTCGAGGGTTGTGCTCCGGTAACCACCACGGCTGTAGAGC 109
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8 GCAGGTAGCAGTGACCCTCGAGGGTTGTGCTCCGGTAACCACCACGGCTGTAGAGC 67

Qy   110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTTCTGAAGA 169
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTTCTGAAGA 127

Qy   170 AGTGGGGTCTCGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128 AGTGGGGTCTCGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 187

Qy   230 TTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCGACTGGCATAGTCGG 289
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 TTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCGACTGGCATAGTCGG 247

Qy   290 CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGACGATGACGGACAA 349

Db   248 CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGACGATGACGGACAA 307
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCTTCTGGGC 409
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCTTCTGGGC 367

Qy   410 AGTTGATGGCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 AGTTGATGGCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCCATATCAT 427

Qy   470 CTACACTGTCAACCTTCACAGCAGCGTTCTCATCCTCGGCCTTCATCAGCCTGGACCGGTA 529
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 CTACACTGTCAACCTTCACAGCAGCGTTCTCATCCTCGGCCTTCATCAGCCTGGACCGGTA 487

Qy   530 CCTCGCCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGAAACTCGTGGCTGAAAAGGC 589
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 CCTCGCCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGAAACTCGTGGCTGAAAAGGC 547

Qy   590 AGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGCCACTATACCTGACTTCATCTTTGC 649
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 AGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGCCACTATACCTGACTTCATCTTTGC 607

Qy   650 CGACGTCAGCCAGGGGGACATCAGTCAGGGGGATGA 685
         ||||||||||||||||||||||||||||||||||||
Db   608 CGACGTCAGCCAGGGGGACATCAGTCAGGGGGATGA 643

RESULT 3
MMLESTRPT
LOCUS   MMLESTRPT   1575 bp   mRNA   linear   ROD 11-NOV-1996
DEFINITION M.musculus mRNA for leukocyte-dervied seven transmembrane domain
            receptor.
ACCESSION   X99582
VERSION   X99582.1   GI:1666648
KEYWORDS   lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Moepps,B., Frodl,R., Keesler,H. and Gierschik,P.
  TITLE   cDNA cloning and genomic organization of a leukocyte-derived seven
            transmembrane domain receptor (LESTR) from mouse: a murine
            homologue of the human HIV-1 entry cofactor fusin
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1575)
  AUTHORS   Moepps,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,
            Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK   Revised by author 11-NOV-96
COMMENT   On Nov 12, 1996 this sequence version replaced gi:1657351.
FEATURES   Location/Qualifiers
     source   1..1575
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="B6CBA"
                  /db_xref="taxon:10090"
                  /haplotype="diploid"
                  /sex="female"
                  /tissue_type="thymus"
                  /dev_stage="6-8 weeks old"
     gene   1..1575
                  /gene="lestr"
     CDS   90..1169
                  /gene="lestr"
                  /codon_start=1
                  /product="leukocyte-derived seven transmembrane domain
                  receptor"
                  /protein_id="CAA67894.1"
                  /db_xref="GI:166649"
                  /db_xref="GOA:P70658"

AUTHORS　　Suzuki,G., Nakata,Y., Uzawa,A., Shirasawa,T., Saito,T. and Mita,K.
TITLE　　　Molecular cloning of murine CXCR-4, a murine homologue of HIV entry
　　　　　　co-receptor, that is expressed on leukocyte and other organs
JOURNAL　　Unpublished
REFERENCE　2　(bases 1 to 1809)
AUTHORS　　Suzuki,G.
TITLE　　　Direct Submission
JOURNAL　　Submitted (29-JAN-1997) Gen Suzuki, National Institute of
　　　　　　Radiological Sciences, Division of Radiation Health; 4-9-1,
　　　　　　Anagawa, Inage-ku, Chiba, Chiba 263, Japan
　　　　　　(E-mail:gsuzuki@unmxl.nirs.go.jp, Tel:043-251-2111,
　　　　　　Fax:043-284-1736)
FEATURES　　　　　　Location/Qualifiers
　　source　　　　　1..1809
　　　　　　　　　　　/organism="Mus musculus"
　　　　　　　　　　　/mol_type="mRNA"
　　　　　　　　　　　/strain="C57BL-6"
　　　　　　　　　　　/db_xref="taxon:10090"
　　　　　　　　　　　/tissue_type="thymus"
　　　　　　　　　　　/dev_stage="E17"
　　CDS　　　　　　　72..1151
　　　　　　　　　　　/note="fusin"
　　　　　　　　　　　/codon_start=1
　　　　　　　　　　　/product="murine CXCR-4"
　　　　　　　　　　　/protein_id="BAA19187.1"
　　　　　　　　　　　/db_xref="GI:1816446"
　　　　　　　　　　　/translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
　　　　　　　　　　　TIYPIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
　　　　　　　　　　　DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVV
　　　　　　　　　　　GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
　　　　　　　　　　　IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
　　　　　　　　　　　ILSKGKRGGHSSVSTESESSSFHSS"
ORIGIN

Query Match　　　　90.3%;　　Score 618.6;　DB 10;　Length 1809;
Best Local Similarity　99.2%;　Pred. No. 2.1e-152;
Matches　632;　Conservative　0;　Mismatches　4;　Indels　1;　Gaps　1;

Qy　　50　GCAGGTAGCAGTG-ACCCTCTGAGCGTTTGGTGTCTCCGGTAACCACCACGGCTGTAGAG　108
　　　　　||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db　　 1　GCAGGTAGCAGTGAAACCTCTGAGCGTTTGGTGTCTCCGGTAACCACCACGGCTGTAGAG　60

Qy　109　CGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAATACTACTCGAAG　168
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　 61　CGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAATACTACTCGAAG　120

Qy　169　AAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCC　228
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　121　AAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCC　180

Qy　229　ATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCG　288
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　181　ATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCG　240

Qy　289　GCAATGGATTGGTGATCCTGGTCAGTGGGTTACCAGGAGAAGCTAAGGAGCATGACGGACA　348
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　241　GCAATGGATTGGTGATCCTGGTCAGTGGGTTACCAGGAGAAGCTAAGGAGCATGACGGACA　300

Qy　349　AGTACCGGCTGCACCTGTCAGTGGCTGACCTGCTGTTCATCACTCTCCCTTCGG　408
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　301　AGTACCGGCTGCACCTGTCAGTGGCTGACCTGCTGTTCATCACTCTCCCTTCGG　360

Qy　409　CAGTTGATGCCATGGCCATGGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCA　468
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　361　CAGTTGATGCCATGGCCATGGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCA　420

Qy　469　TCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGT　528
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　421　TCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGT　480

Qy　529　ACCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGG　588
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

（左段 — 回転ページ下部）

/db_xref="SWISS-PROT:P70658"
/translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
TIYPIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVVV
GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
ILSKGKRGGHSSVSTESESSSFHSS"
polyA_signal　　1498..1503
　　　　　　　　　/gene="lestr"
ORIGIN

Query Match　　　　92.1%;　　Score 631.2;　DB 10;　Length 1575;
Best Local Similarity　99.5%;　Pred. No. 9.6e-156;
Matches　633;　Conservative　0;　Mismatches　3;　Indels　0;　Gaps　0;

Qy　　50　GCAGGTAGCAGTGACCCTCTGAGCGTTTGGTGTCTCCGGTAACCACCACGGCTGTAGAGC　109
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　20　GCAGGTAGCAGTGACCCTCTGAGCGTTTGGTGTCTCCGGTAACCACCACGGCTGTAGAGC　79

Qy　110　GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAATACTACTCGAAGA　169
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　80　GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAATACTACTCGAAGA　139

Qy　170　AGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA　229
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　140　AGTGGGGTTCTGAGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA　199

Qy　230　TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGG　289
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　200　TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGG　259

Qy　290　CAATGGATTGGTGATCCTGGTCAGTGGGTTACCAGGAGAAGCTAAGGAGCATGACGGACA　349
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　260　CAATGGATTGGTGATCCTGGTCAGTGGGTTACCAGGAGAAGCTAAGGAGCATGACGGACA　319

Qy　350　GTACCGGCTGCACCTGTCAGTGGCTGACCTGCTGTTCATCACTCTCCCTTCGGGGC　409
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　320　GTACCGGCTGCACCTGTCAGTGGCTGACCTGCTGTTCATCACTCTCCCTTCGGGGC　379

Qy　410　AGTTGATGCCATGGCCATGGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCAT　469
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　380　AGTTGATGCCATGGCCATGGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCAT　439

Qy　470　CTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTA　529
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　440　CTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGACCGGTA　499

Qy　530　CCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGC　589
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　500　CCTGCCATTGTCCACGCCACCAACAGTCAGAGGCCAAGGAAACTGCTGGCTGAAAAGGC　559

Qy　590　AGTCTATGTGGGGCGTTGGATCCCCAGCCCTCCTGACTATACCCTGACTTCATCTTTGC　649
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　560　AGTCTATGTGGGGCGTTGGATCCCCAGCCCTCCTGACTATACCCTGACTTCATCTTTGC　619

Qy　650　CGACGTCAGCCAGGGGACATCAGTCAGGGGGATGA　685
　　　　　|||||||||||||||||||||||||||||||||||
Db　620　CGACGTCAGCCAGGGGACATCAGTCAGGGGGATGA　655

```
Db    481 ACCTCGCTATTGTCCACGCCACCAACAGTCCAGGAGGCCAAGGAAACTGTCGGCTGAAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    589 CAGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGCTGACTATACCTGACTTCATCTTTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGCTGACTATACCTGACTTCATCTTTG 600

Qy    649 CCGACGTCAGCCAGGGGACATCAGTCAGGGGATGA 685
              ||||||||||||||||||||||||||||||||||
Db    601 CCGACGTCAGCCAGGGGACATCAGTCAGGGGATGA 637
```

RESULT 5
MMU9760
LOCUS   MMU59760   1223 bp   mRNA   linear   ROD 30-SEP-2002
DEFINITION Mus musculus fusin homolog mRNA, complete cds.
ACCESSION   U59760
VERSION   U59760.1 GI:1527134
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1223)
  AUTHORS   Heesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
  TITLE   Cloning of the mouse fusin gene, homologue to a human HIV-1
            co-factor
  JOURNAL   J. Immunol. 157 (12), 5455-5460 (1996)
  MEDLINE   97113334
  PUBMED   8955194
REFERENCE   2 (bases 1 to 1223)
  AUTHORS   Heesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1996) Pathology, Harvard Medical School, 200
            Longwood Avenue, Boston, MA 02115, USA
FEATURES   Location/Qualifiers
     source   1..1223
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129/Sv"
                     /db_xref="taxon:10090"
     CDS   30..1109
                     /note="seven transmembrane spanning receptor;
                     chemokine-receptor like"
                     /codon_start=1
                     /product="fusin homolog"
                     /protein_id="AAB07725.1"
                     /db_xref="GI:1527135"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHAINSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDRYICDRLYPDSLWMVVFQFQHIMVGLIILP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMRGSSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"

ORIGIN

Query Match   86.9%; Score 595; DB 10; Length 1223;
Best Local Similarity 100.0%; Pred. No. 3.5e-146;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     91 ACCACCACCGGCTGTAGAGCCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTT 150
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACCACCGGCTGTAGAGCCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTT 60

Qy    151 CTGATAACTACTCTGAAGAGTGGGGTCTGGAGACTATGATGACTCCAACAGGAACCCTGCT 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGATAACTACTCTGAAGAGTGGGGTCTGGAGACTATGATGACTCCAACAGGAACCCTGCT 120

Qy    211 TCCGGGATGAAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCT 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCCGGGATGAAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCT 180

Qy    271 TCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTCGTGATGGGTTACCAGAAGAAGC 330
```

```
Db    181 TCTTGACTGGCATAGTCGGCAATGATGGTGATCCTCGTGATGGGTTCATGGGTTACCAGAAGAAGC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    331 TAAGGAGCATGACGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCA 390
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TAAGGAGCATGACGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCA 300

Qy    391 TCACACTCCCTTCTGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGAATTTTTGT 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCACACTCCCTTCTGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGAATTTTTGT 360

Qy    451 GTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGCCT 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGCCT 420

Qy    511 TCATCAGCCTGGACCGGTACCTCGCCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGA 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCATCAGCCTGGACCGGTACCTCGCCATTGTCCAGCCACCAACAGTCAAAGGCCAAGGA 480

Qy    571 AACTGCTGCGGTCTATGTGGGGTCTGATCCCAGCGCCTCCTCCTGACTA 630
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACTGCTGCGGTCTATGTGGGGTCTGATCCCAGCGCCTCCTCCTGACTA 540

Qy    631 TACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGGATGA 685
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGGATGA 595
```

RESULT 6
MMLCR12
LOCUS   MMLCR12   1180 bp   mRNA   linear   ROD 18-NOV-1997
DEFINITION Mus musculus lcr-1 gene.
ACCESSION   Z80112
VERSION   Z80112.1 GI:2632100
KEYWORDS   CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1180)
  AUTHORS   Schubel,A., Burgstahler.R. and Lipp,M.
  TITLE   The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
            High expression in thymus and lymphoid tissues
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1180)
  AUTHORS   Lipp,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
            Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
            D-13122, GERMANY
COMMENT   On Nov 21, 1997 this sequence version replaced gi:1542888.
FEATURES   Location/Qualifiers
     source   1..1180
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129SV"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
                     /dev_stage="adult"
     gene   1..1180
                     /gene="lcr-1"
     CDS   1..1074
                     /gene="lcr-1"
                     /function="receptor for chemokine SDF-1"
                     /standard_name="CXC chemokine receptor 4"
                     /codon_start=1
                     /product="CXCR-4"
                     /protein_id="CAB02202.1"
                     /db_xref="GI:1542889"
                     /db_xref="GOA:P70658"
                     /db_xref="SWISS-PROT:P70658"
                     /translation="MEPISIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTI
                     YFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADW

FEATURES   Location/Qualifiers
     source   1..1422
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV"
                     /db_xref="taxon:10090"
                     /clone_lib="lambda fix (TM) II"
                     /dev_stage="adult"
     intron   <1..259
                     /number=1
     gene   260..1422
                     /gene="lcr-1"
     exon   260..>1422
                     /gene="lcr-1"
                     /product="CXCR-4"
                     /standard_name="CXC chemokine receptor 4"
                     /number=2
                     /function="receptor for chemokine SDF-1"

ORIGIN

Query Match   79.7%; Score 545.8; DB 10; Length 1422;
Best Local Similarity   97.9%; Pred. No. 3.6e-133;
Matches 553; Conservative   0; Mismatches 12; Indels   0; Gaps   0;

Qy   121 TGGAACGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG 180
          |||| ||||||||||||||||||||||||||||||||||||||
Db   240 TGTCCCTGTTTCCTTTGCAGATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG 299

Qy   181 GAGACTATGACTCCAACAACAAGGAACCCGCTTCCGGGATGAAAAGTCCATTCAATAGGA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   300 GAGACTATGACTCCAACAACAAGGAACCCGCTTCCGGGATGAAAAGTCCATTCAATAGGA 359

Qy   241 TCTTCCTGCCCACCATCTACTTCTTGATCGGCATGTCGACTGCATGTCGGCAATGATTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 TCTTCCTGCCCACCATCTACTTCTTGATCGGCATGTCGACTGCATGTCGGCAATGATTGG 419

Qy   301 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGACTAAGGACATGACGGACAAGCTGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   420 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGACTAAGGACATGACGGACAAGCTGC 479

Qy   361 ACCTGCAGTGGCTGACCTCCCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGATGCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   480 ACCTGCAGTGGCTGACCTCCCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGATGCCA 539

Qy   421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCTACACTGTCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   540 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCTACACTGTCA 599

Qy   481 ACCTTCACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTTCATCAGCCTTCATGTGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600 ACCTTCACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTTCATCAGCCTTCATGTGG 659

Qy   541 TCCACGCCACCAACAGTCAAAGGCAAGGAAACTGCTGGCAGAAACTGCTGGCTGAAAAG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660 TCCACGCCACCAACAGTCAAAGGCAAGGAAACTGCTGGCAGAAACTGCTGGCTGAAAAG 719

Qy   601 GCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720 GCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 779

Qy   661 AGGGGGACATCAGTCAGGGGGATGA 685
          |||||||||||||||||||||||||
Db   780 AGGGGGACATCAGTCAGGGGGATGA 804

---

exon   YFGKFLCKAVHIIYTVNLYSSVLLIAFISLDRYLAIVHATNSQRPRKLLAEKAVYVGV
              WIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLMVVFQFQHIMVGLILPGI
              VILSCYCIIISKLSHSKGHQKRKALKTTVILLAFFACWLPYYVGISIDSFILLGVIK
              QGCDFESIVHKWISITEALAFFHCCLNPIILYAFLGAKFKSSAQHALNSMSRGSSLKIL
              SKGKRGGHSSVSTESESSSFHSS"
exon   <1..15
              /gene="lcr-1"
              /number=1
exon   16..>1180
              /gene="lcr-1"
              /number=2

ORIGIN

Query Match   80.1%; Score 548.4; DB 10; Length 1180;
Best Local Similarity   99.8%; Pred. No. 7.3e-134;
Matches 549; Conservative   0; Mismatches 1; Indels   0; Gaps   0;

Qy   136 TGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCA 195
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 TCAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCA 70

Qy   196 ACAAGGAACCCGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCA 255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 ACAAGGAACCCGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCA 130

Qy   256 TCTACTTCATCATCTTCTTGACTGGCATAGTCGGCATAGTGGATTGGTGATCCTGGTCATGG 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 TCTACTTCATCATCTTCTTGACTGGCATAGTCGGCATAGTGGATTGGTGATCCTGGTCATGG 190

Qy   316 GTTACCAGAAGAAGCTAAGGACGATGACGGACAAGCTGCACCTGTCAGTGGCTG 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 GTTACCAGAAGAAGCTAAGGACGATGACGGACAAGCTGCACCTGTCAGTGGCTG 250

Qy   376 ACCTCCCTCTTTGTCATCACACTCCCCTTCTCGGCAGTTGATGCCATGGCTGACTGGTACT 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 ACCTCCCTCTTTGTCATCACACTCCCCTTCTCGGCAGTTGATGCCATGGCTGACTGGTACT 310

Qy   436 TTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCTACACTGTCAACCTTCACAGCAGCG 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 TTGGGAAATTTTGTGTAAGGCTGTCCATATCATCATCTACACTGTCAACCTTCACAGCAGCG 370

Qy   496 TTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCCGGTGAAAAGCCAAGGACCACCAACA 555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 TTCTCATCCTGGCCTTCATCAGCCTGACCGGTACCCGGTGAAAAGCCAAGGACCACCAACA 430

Qy   556 GTCAAAGGCCAAGGAAACTGCTGGCTATGTGGGGCTGTCTATGTGGGGCTGTCTGGATCCCAG 615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 GTCAAAGGCCAAGGAAACTGCTGGCTATGTGGGGCTGTCTATGTGGGGCTGTCTGGATCCCAG 490

Qy   616 CCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTC 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 CCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTC 550

Qy   676 AGGGGGATGA 685
          ||||||||||
Db   551 AGGGGGATGA 560

RESULT 7
MMLCR13

RESULT 8

Qy   241 TCTTCCTGCCCACCATCTACTTCATCATCTTCTGACTGGCATAGTCGGCAATGGATTGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2408 TCTTCCTGCCCACCATCTACTTCATCATCTTCTGACTGGCATAGTCGGCAATGGATTGG 2467

Qy   301 TGATCCTGGTCATGGGTACCAGAGAAGCTAGGAGCATGAGGAGGACGGACAAGTACGGCTGC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2468 TGATCCTGGTCATGGGTACCAGAGAAGCTAGGAGCATGAGGAGGACGGACAAGTACGGCTGC 2527

Qy   361 ACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTGGGCCAGTTGATGCCA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2528 ACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTGGGCCAGTTGATGCCA 2587

Qy   421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2588 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 2647

Qy   481 ACCTCACAGCAGCGGTTCTCATCCTGCCTTCATCAGCCTGACCGTACCTCGCCATTG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2648 ACCCTCACAGCAGCGGTTCTCATCCTGCCTTCATCAGCCTGACCGTACCTCGCCATTG 2707

Qy   541 TCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2708 TCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 2767

Qy   601 GCGTCTGATCCCAGCCTCCTCGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2768 GCGTCTGGATCCCAGCCTCCTCGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 2827

Qy   661 AGGGGGACATCAGTCAGGGGGATGA 685
          |||||||||||||||||||||||||
Db   2828 AGGGGGACATCAGTCAGGGGGATGA 2852

RESULT 9

LOCUS   MMLESTRGN   3770 bp   DNA   linear   ROD 02-APR-1997
DEFINITION M.musculus gene encoding leukocyte-derived seven transmembrane
            domain receptor, strain B6.
ACCESSION   X99581
VERSION   X99581.1 GI:1924959
KEYWORDS   lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Moepps,B., Frodl,R., Kessler,H. and Gierschik,P.
  TITLE   cDNA cloning and genomic organization of a leukocyte-derived seven
            transmembrane domain receptor (LESTR) from mouse: a murine
            homologue of the human HIV-1 entry cofactor fusin
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Moepps,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,
            Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
  REMARK   Revised by [3]
REFERENCE   3 (bases 1 to 3770)
  AUTHORS   Moepps,B.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-APR-1997) B. Moepps, Universitaet Ulm,
            Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
COMMENT   On Apr 3, 1997 this sequence version replaced gi:1666646.
FEATURES   Location/Qualifiers
     source   1..3770
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="B6"
                     /db_xref="taxon:10090"
                     /haplotype="diploid"
                     /tissue_type="thymus"
                     /dev_stage="6-8 weeks old"
     exon   <1..110

LOCUS   MMU65580   3366 bp   DNA   linear   ROD 13-DEC-1996
DEFINITION Mus musculus fusin (CXCR-4) gene, complete cds.
ACCESSION   U65580
VERSION   U65580.1 GI:1731650
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 3366)
  AUTHORS   Heesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
  TITLE   Cloning of the mouse fusin gene, homologue to a human HIV-1
            co-factor
  JOURNAL   J. Immunol. 157 (12), 5455-5460 (1996)
  MEDLINE   97113334
  PUBMED   8955194
REFERENCE   2 (bases 1 to 3366)
  AUTHORS   Heesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
  TITLE   Direct Submission
  JOURNAL   Submitted (30-JUL-1996) Pathology, Harvard Medical School, 200
            Longwood Avenue, Boston, MA 02115, USA
FEATURES   Location/Qualifiers
     source   1..3366
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
     gene   1..3366
                     /gene="CXCR-4"
     CDS   join(1..21,2308..3366)
                     /gene="CXCR-4"
                     /note="SDF-1 chemokine receptor"
                     /codon_start=1
                     /product="fusin"
                     /protein_id="AAC52953.1"
                     /db_xref="GI:1645915"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIPLP
                     TIYFIFLTGIVGNGLVLVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWFGKFLCKAVHIIYTVNLYSSVLILAPISLDRYLAIVHATNSQPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                     GIVILSVCYIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"
     exon   1..21
                     /gene="CXCR-4"
                     /number=1
     exon   1..15
                     /gene="CXCR-4"
                     /note="alternate exon 1"
     intron   16..>2307
                     /gene="CXCR-4"
                     /note="alternate intron 1"
     intron   22..2307
                     /gene="CXCR-4"
                     /number=1
     exon   2308..3366
                     /gene="CXCR-4"
                     /number=2

ORIGIN

Query Match   79.7%;   Score 545.8;   DB 10;   Length 3366;
Best Local Similarity   97.9%;   Pred. No. 3.7e-133;
Matches 553;   Conservative 0;   Mismatches 12;   Indels 0;   Gaps 0;

Qy   121 TGGAACCGATCAGTCGAGTGTGAGTATATACACTTCTGATAACTACACTTCTGAAGAAGTGGGGTCTG 180
          ||||| |||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   2288 TGTCCCTGTTTCCTTTGCAGATATACACTTCTGATAACTACACTACTCTGAAGAAGTGGGGTCTG 2347

Qy   181 GAGACTATGACTCCAACAAGGAAACCCTGCTTCCCGGGATGAAAACGTCCATTTCAATAGGA 240
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   2348 GAGACTATGACTCCAACAACGAACCCCTGCTTCCCGGGATGAAAAACGTCCATTTCAATAGGA 2407

```
gene        /number=1
            join(90..110,2306..3364)
            /gene="lestr"
CDS         join(90..110,2306..3364)
            /gene="lestr"
            /codon_start=1
            /product="leukocyte-derived seven transmembrane domain
            receptor"
            /protein_id="CAA67893.1"
            /db_xref="GI:1666647"
            /db_xref="GOA:P70658"
            /db_xref="SWISS-PROT:P70658"
            /translation="MEPISVSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLP
            TIYFIFLFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
            DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
            GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVFQFQHIMVGLVLP
            GIVILSCYCIIISKLSHSKGHQKRKALKTTVILLLAFFACWLPYYVGISIDSFILLGV
            IKQGGDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
            ILSKGKRGGHSSVSTESESSFHSS"
misc_feature  105..2305
              /gene="lestr"
              /note="splice variant/intron"
intron      111..2305
            /gene="lestr"
            /number=1
exon        2306..>3770
            /gene="lestr"
            /number=2
polyA_signal  3693..3698
ORIGIN
```

Query Match   79.0%;   Score 541;   DB 10;   Length 3770;
Best Local Similarity   97.3%;   Pred. No. 7e-132;
Matches 550;   Conservative 0;   Mismatches 15;   Indels 0;   Gaps 0;

```
Qy    121 TGGAACCGATCAGTGTGAGTATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG 180
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   2286 TGTCCCTGTTTCCTTTGCAGATATACACTTCTGATAACTACTCTGAAGAAGTGGGTTCTG 2345

Qy    181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2346 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 2405

Qy    241 TCTTCCTGCCCACCATCTACTTCATCTTCTTGACTGGCATGGTCGGCAATGGATTGG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2406 TCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATGGTCGGCAATGGATTGG 2465

Qy    301 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2466 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGC 2525

Qy    361 ACCTGTCAGTGGCTGACCTCCTGTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2526 ACCTGTCAGTGGCTGACCTCCTGTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCA 2585

Qy    421 TGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCATATCATCTACACTGTCA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2586 TGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCATATCATCTACACTGTCA 2645

Qy    481 ACCTCTACAGCAGCAGGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGGCCATTG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2646 ACCTCTACAGCAGCAGCAGGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGGCCATTG 2705

Qy    541 TCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2706 TCCACGCCACCAACAGTCAGAGGCCCAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 2765

Qy    601 GCGTCTGGATCCCAGCGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2766 GCGTCTGGATCCCAGCCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 2825

Qy    661 AGGGGGACATCAGTCAGGGGGATGA 685
          |||||||||||||||||||||||||
Db   2826 AGGGGGACATCAGTCAGGGGGATGA 2850
```

```
                79965    80064: gap of unknown length
        *       80065    81745: contig of 1681 bp in length
        *       81746    81845: gap of unknown length
        *       81846    83417: contig of 1572 bp in length
        *       83418    83517: gap of unknown length
        *       83518    84910: contig of 1393 bp in length
        *       84911    85010: gap of unknown length
        *       85011    86788: contig of 1778 bp in length
        *       86789    86888: gap of unknown length
        *       86889    88892: contig of 2004 bp in length
        *       88893    88992: gap of unknown length
        *       88993    90670: contig of 1678 bp in length
        *       90671    90770: gap of unknown length
        *       90771    93207: contig of 2437 bp in length
        *       93208    93307: gap of unknown length
        *       93308    96658: contig of 3351 bp in length
        *       96659    96758: gap of unknown length
        *       96759   100576: contig of 3818 bp in length
        *      100577   100676: gap of unknown length
        *      100677   104988: contig of 4312 bp in length
        *      104989   105088: gap of unknown length
        *      105089   108893: contig of 3805 bp in length
        *      108894   108993: gap of unknown length
        *      108994   112632: contig of 3639 bp in length
        *      112633   112732: gap of unknown length
        *      112733   116328: contig of 3596 bp in length
        *      116329   116428: gap of unknown length
        *      116429   120980: contig of 4552 bp in length
        *      120981   121080: gap of unknown length
        *      121081   126821: contig of 5741 bp in length
        *      126822   126921: gap of unknown length
        *      126922   131331: contig of 4410 bp in length
        *      131332   131431: gap of unknown length
        *      131432   138705: contig of 7274 bp in length
        *      138706   138805: gap of unknown length
        *      138806   145169: contig of 6364 bp in length
        *      145170   145269: gap of unknown length
        *      145270   150696: contig of 5427 bp in length
        *      150697   150796: gap of unknown length
        *      150797   159988: contig of 9192 bp in length
        *      159989   160088: gap of unknown length
        *      160089   167244: contig of 7156 bp in length
        *      167245   167344: gap of unknown length
        *      167345   175370: contig of 8026 bp in length.

FEATURES             Location/Qualifiers
     source          1..175370
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="UNK"
                     /clone="RP43-21D9"
     misc_feature    1..5345
                     /note="assembly_name:Contig100"
     misc_feature    5446..17195
                     /note="assembly_name:Contig101"
     misc_feature    17296..33250
                     /note="assembly_name:Contig102"
     misc_feature    33351..45646
                     /note="assembly_name:Contig103"
     misc_feature    45747..61992
                     /note="assembly_name:Contig104"
     misc_feature    62093..63423
                     /note="assembly_name:Contig16"
     misc_feature    63524..64666
                     /note="assembly_name:Contig93"
     misc_feature    64767..65994
                     /note="assembly_name:Contig45"
     misc_feature    66095..67430
                     /note="assembly_name:Contig48"
     misc_feature    67551..68982
                     /note="assembly_name:Contig5"
     misc_feature    69083..70891
                     /note="assembly_name:Contig56"
```

```
     misc_feature    70992..72172
                     /note="assembly_name:Contig60"
     misc_feature    72273..73614
                     /note="assembly_name:Contig67"
     misc_feature    73715..75185
                     /note="assembly_name:Contig68"
     misc_feature    75286..76854
                     /note="assembly_name:Contig69"
     misc_feature    76955..78413
                     /note="assembly_name:Contig73"
     misc_feature    78514..79964
                     /note="assembly_name:Contig75"
     misc_feature    80065..81745
                     /note="assembly_name:Contig77"
     misc_feature    81846..83417
                     /note="assembly_name:Contig79"
     misc_feature    83518..84910
                     /note="assembly_name:Contig80"
     misc_feature    85011..86788
                     /note="assembly_name:Contig81"
     misc_feature    86889..88892
                     /note="assembly_name:Contig82"
     misc_feature    88993..90670
                     /note="assembly_name:Contig83"
     misc_feature    90771..93207
                     /note="assembly_name:Contig84"
     misc_feature    93308..96658
                     /note="assembly_name:Contig85"
     misc_feature    96759..100576
                     /note="assembly_name:Contig86"
     misc_feature    100677..104988
                     /note="assembly_name:Contig87"
     misc_feature    105089..108893
                     /note="assembly_name:Contig88"
     misc_feature    108994..112632
                     /note="assembly_name:Contig89"
     misc_feature    112733..116328
                     /note="assembly_name:Contig90"
     misc_feature    116429..120980
                     /note="assembly_name:Contig91"
     misc_feature    121081..126821
                     /note="assembly_name:Contig92"
     misc_feature    126922..131331
                     /note="assembly_name:Contig93"
     misc_feature    131432..138705
                     /note="assembly_name:Contig94"
     misc_feature    138806..145169
                     /note="assembly_name:Contig95"
     misc_feature    145270..150696
                     /note="assembly_name:Contig96"
     misc_feature    150797..159988
                     /note="assembly_name:Contig97"
     misc_feature    160089..167244
                     /note="assembly_name:Contig98"
     misc_feature    167345..175370
                     /note="assembly_name:Contig99"

ORIGIN

Query Match           79.0%;  Score 541;  DB 2;  Length 175370;
Best Local Similarity 97.3%;  Pred. No. 8.5e-132;
Matches 550;  Conservative  0;  Mismatches 15;  Indels  0;  Gaps  0;

Qy      121 TGGAACCGATCAGTGTGAGTATATACACTCTGATAACTACTCTGAGAAGTGGGTCTG 180
                |||||||||||||||| |||||||||||| ||||||||||||||||||||||| ||||
Db   167962 TGTCCTGTTTCCTTTGCAGATATACACTTCTGATAACTACTCTGAGAAGTGGGTTCTG 168021

Qy      181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   168022 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 168081

Qy      241 TCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 300
                |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   168082 TCTTCCTGCCACCATCTACTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 168141
```

```
FEATURES             Location/Qualifiers
     source          1..185215
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP23-480N14"
     misc_feature    1..1119
                     /note="assembly_name:Contig51"
     misc_feature    1220..2637
                     /note="assembly_name:Contig103"
     misc_feature    2738..5059
                     /note="assembly_name:Contig122"
     misc_feature    5160..6203
                     /note="assembly_name:Contig124"
     misc_feature    6304..11814
                     /note="assembly_name:Contig126"
     misc_feature    11915..18453
                     /note="assembly_name:Contig127"
     misc_feature    18554..33684
                     /note="assembly_name:Contig128"
     misc_feature    33785..57713
                     /note="assembly_name:Contig129"
     misc_feature    57814..84551
                     /note="assembly_name:Contig130"
     misc_feature    84652..122285
                     /note="assembly_name:Contig131"
     misc_feature    122386..185215
                     /note="assembly_name:Contig132"
ORIGIN

Query Match       79.0%;  Score 541;  DB 2;  Length 185215;
Best Local Similarity   97.3%;  Pred. No. 8.5e-132;
Matches  550;  Conservative   0;  Mismatches  15;  Indels   0;  Gaps   0;

Qy        121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTG 180
                  ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91301 TGTCCCTGTTCCTTGCAGATATACACTTCTGATAACTACTCTGAAGAAGTGGGTTCTG 91360

Qy        181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91361 GAGACTATGACTCCAACAAGGAACCCTGTTCCGGATGAAAACGTCCATTCAATAGA 91420
```

```
Qy        301 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACCAAGTACCGGCTGC 360
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     168142 TGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACCAAGTACCGGCTGC 168201

Qy        361 ACCTGTCAGTGGCTGACCTCCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168202 ACCTGTCAGTGGCTGACCTCCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCA 168261

Qy        421 TGGCTGACTGGTTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCAATCATCTACACTGTCA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168262 TGGCTGACTGGTTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 168321

Qy        481 ACCTCTACAGCAGCGGTTCTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGGCCATTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168322 ACCTCTACAGCAGCGGTTCTCATCCTGCCTTCATCAGCCTGGACCGGTACCTCGGCTATTG 168381

Qy        541 TCCACGCCACCAACAGTCAAAGGCCAAGGAAACTCTGGCTGAAAAGGCAGTCTATGTGG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168382 TCCACGCCACCAACAGTCAGAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGG 168441

Qy        601 GCGTCTGGATCCCAGCCCTCCCTGCTATACCTGACTTCATCTTTGCCGACGTCAGCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168442 GCGTCTGGATCCCAGCCCTCCCTGCTATACCTGACTTCATCTTTGCCGACGTCAGCC 168501

Qy        661 AGGGGGACATCAGTCAGGGGATGA 685
              ||||||||||||||||||||||||
Db     168502 AGGGGGACATCAGTCAGGGGGGATGA 168526
```

RESULT 12

```
Qy    241  TCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  91421  TCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG  91480

Qy    301  TGATCCTGGTCATGGGTTACCAGAAGAGCTAGGAGCATGACGGACAAGTACCGGCTGC   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  91481  TGATCCTGGTCATGGGTTACCAGAAGAGCTAGGAGCATGACGGACAAGTACCGGCTGC   91540

Qy    361  ACCTGCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCGGGCAGTTGATGCCA   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  91541  ACCTGCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCGGGCAGTTGATGCCA   91600

Qy    421  TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  91601  TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA   91660

Qy    481  ACCTCTACAGCAGCAGCGTTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCCATTG   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  91661  ACCTCTACAGCAGCAGCGTTCATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCTATTG   91720

Qy    541  TCCAGCCACCACACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAGGCAGTCTATGTGG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  91721  TCCAGCCACCACACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAGGCAGTCTATGTGG   91780

Qy    601  GCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  91781  GCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCC   91840

Qy    661  AGGGGGACATCAGTCAGGGGGATGA   685
           |||||||||||||||||||||||||
Db  91841  AGGGGGACATCAGTCAGGGGGATGA   91865
```

RESULT 12

```
MMLCR11
LOCUS       MMLCR11                    528 bp      mRNA    linear    ROD 07-OCT-1996
DEFINITION  Mus musculus mRNA encoding CXCR-4.
ACCESSION   Z80111
VERSION     Z80111.1    GI:1542886
KEYWORDS    CXCR-4.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Schubel,A., Burgstahler,R. and Lipp,M.
  TITLE     The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
            High expression in thymus and lymphoid tissues
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 528)
  AUTHORS   Lipp,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
            Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
            D-13122, GERMANY
FEATURES             Location/Qualifiers
     source          1..528
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129SV"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
                     /dev_stage="adult"
     exon            <1..98
                     /number=1
     gene            join(84..98,99..528)
                     /gene="lcr-1"
     CDS             join(84..98,99..>528)
                     /gene="lcr-1"
                     /function="receptor for chemokine SDF-1"
                     /standard_name="CXC chemokine receptor 4"
                     /codon_start=1
                     /product="CXCR-4"
```

```
                     /protein_id="CAB02201.1"
                     /db_xref="GI:1542887"
                     /db_xref="GOA:P70658"
                     /db_xref="SWISS-PROT:P70658"
                     /translation="MEPISIYTSDNYSEVGSGYDYSNKEPCFRDENVHFNRIFLPTI
                     YFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVADW
                     YFGKFLCKAVHIIYTVNLYSSVLLLAFISLDRYLAIVHATNSQRPR"
     exon            99..>528
                     /gene="lcr-1"
                     /number=2
ORIGIN
```

```
Query Match      72.4%;    Score 495.8;   DB 10;   Length 528;
Best Local Similarity   98.5%;   Pred. No. 5.6e-120;
Matches  513;  Conservative  0;   Mismatches  2;   Indels  6;   Gaps  1;
```

```
Qy     50  GCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC  109
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     14  GCAGGTAGCAGTGACCCTCTGAGGCGTTGGGCGTTGGGGCCTCGGGTAACCACCACGGCTGTAGAGC   73

Qy    110  GAGTGGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCGATAAACTACTCTGAAGA  169
           |||||||||||||||||||||||||||       |||||||||||||||||||||||||
Db     74  GAGTGTTGCCATGGAACCGATC------AGTATATACACTTCTGATAACTACTCTGAAGA  127

Qy    170  AGTGGGGTCTGGAGACTATGACTCCAACAACAGGAACCCTGCTTCCGGATGAAAAGTCCA  229
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128  AGTGGGGTCTGGAGACTATGACTCCAACAACAGGAACCCTGCTTCCGGGATGAAAAGTCCA  187

Qy    230  TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGG  289
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188  TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGG  247

Qy    290  CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAA  349
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248  CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAA  307

Qy    350  GTACCGGCTGCACCTGCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCGGGC  409
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308  GTACCGGCTGCACCTGCAGTGGCTGACCTCCTTGTCATCACACTCCCCTTCGGGC  367

Qy    410  AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT  469
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368  AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT  427

Qy    470  CTACACTGTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTA  529
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428  CTACACTGTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTA  487

Qy    530  CCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGA  570
           |||||||||||||||||||||||||||||||||||||||||
Db    488  CCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGA  528
```

RESULT 13

```
AC122097
LOCUS       AC122097               190149 bp    DNA    linear    HTG 15-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-108A24, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC122097
VERSION     AC122097.4  GI:25008263
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 190149)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,A., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
```

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,R., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Haml,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladin,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,K., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montenayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelemeh,O., Okwuonu,G., Olampunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,J., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,A., Tabor,P., Taylor,C.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Wilson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhou,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 190149)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (22-MAY-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 190149)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 15, 2002 this sequence version replaced gi:23269619.
          The sequence in this assembly is a combination of BAC based reads
          and whole genome shotgun sequencing reads assembled using Atlas
          (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
          in the feature table below represents a scaffold in the Atlas
          assembly (a 'contig-scaffold'). Within each contig-scaffold,
          individual sequence contigs are ordered and oriented, and separated
          by sized gaps filled with Ns to the estimated size. The sequence
          may extend beyond the ends of the clone and there may be sequence
          contigs within a contig-scaffold that consist entirely of whole
          genome shotgun sequence reads. Both end sequences and whole genome
          shotgun sequence only contigs will be indicated in the feature
          table.
          ------- Genome Center
          Center: Baylor College of Medicine
          Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
------- Project Information -------
Center project name: GXCR
Center clone name: CH230-108A24
------- Summary Statistics -------
Assembly program: Phrap; version 0.990329
Consensus quality: 179135 bases at least Q40
Consensus quality: 181484 bases at least Q30
Consensus quality: 183011 bases at least Q20
Estimated insert size: 187231; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*   1   40067: contig of 40067 bp in length
*   40068   40167: gap of unknown length
*   40168 114460: contig of 74493 bp in length
* 114461 114760: gap of unknown length
* 114761 188186: contig of 73426 bp in length
* 188187 188286: gap of unknown length
* 188287 190149: contig of 1863 bp in length.

FEATURES   Location/Qualifiers
     source   1..190149
                  /organism="Rattus norvegicus"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:10116"
                  /clone="CH230-108A24"

ORIGIN

Query Match   69.1%; Score 473.4; DB 2; Length 190149;
Best Local Similarity 90.8%; Pred. No. 6.3e-114;
Matches 504; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

```
Qy    119 CATGGAACCGATCAGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTC 178
          ||||||||| ||| ||| |||||||||||||| |||||||||||  ||||||||||||
Db  20640 CCTGTCCCTGTTCTTTTGCAGATATACACTTCGGATAACTACTTCCGAAGAAGTAGGGGTC 20699

Qy    179 TGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTTCAATAG 238
          ||||||||||||||||||||||||||||||||||||||||||||||    |||||||||
Db  20700 TGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGAAAACTTCAACAG 20759

Qy    239 GATCTTCCTGCCCACCATCTACTTCATCATTCTTGACTGGCATAGTCGGCAATGGATT 298
          ||||||||||||||||||||   | ||| || |||||||||||||  ||||||||  |
Db  20760 GATCTTCCTGCCCACCATCATTTATCATCTTCTTGACTGGCATAGTGGGCAATGGGT 20819

Qy    299 GGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCT 358
          ||| ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  20820 GGTAATCCTGGTCATGGGTTACCAGAAGAAGCTGAGGAGCATGACGGACAAGTACCGGCT 20879

Qy    359 GCACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTGGCCAGTTGATGC 418
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||| |  ||
Db  20880 GCACCTGTCCGTGGCTGACCTCCTTTGTCATCACACTCCCCTTCTGGCCAGTCGGACGC 20939

Qy    419 CATGGCTGACTTGGGAAATTTTTGTGTAGAGGCTGTCCATATCATCACACTGT 478
          ||||||||||||||||||||||||||||  ||||||| || || ||||||||||| |||
Db  20940 CATGGCTGACTTGGGAAATTTTATGTAGGCTGTGCATATCATCATCACACCGT 20999

Qy    479 CAACCTTCACAGAGCGTTCATCCTGGCCTTCATCAGCCTGGACCGGTTCATCCGCCAT 538
          ||||||| |||| |||||||||||||||||||||||||||||||||||| ||  || ||
Db  21000 CAACCTTTACAGCAGTGTTCATCCTGGCCTTCATCAGCCTGGACCGGTTCACCCTGCCAT 21059

Qy    539 TGTCCAGCGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGT 598
          ||||||||||||||||||| || ||||| ||||| |||||||||||||||||||| ||||||
Db  21060 TGTCCAGCGCCACCAACAGCCAGGAGGCCGGAGGAAGCTGCTGGCTGAAAAGGCCGTCTATGT 21119
```

```
Qy   256  TCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTCAATCCTGGTCATGG  315
          ||||||| |||||||| ||||||||||||||||||| |||||||||| ||||||||||||
Db   122  TCTATTTTATCATCTTCTTGACTGGCATAGTGGGCAATAGTGGCAATGGGTTGGTAATCCTGGTCATGG  181

Qy   316  GTTACCAGAAGCTAAGGACATAGCGGACAAGTACCGGCTGCACCTGTCAGTGGCTG  375
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   182  GTTACCAGAGAGCTGAGGACATGAGCATGACAGACAAGTACCGGCTGCACCTGTCCGTGGCTG  241

Qy   376  ACCTCCTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGGTACT  435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242  ACCTCCTTGTCATCACACTCCCCTTCTGGGCAGTGATCATCACACTGGCTGGTACT  301

Qy   436  TTGGAAATTTTGTTAAGGCTGTCCATATCATCATCTACCAGTGTCAACCTCACAGCAGCG  495
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   302  TTGGAAATTTTATGTGTGCATATCAGTAAGGCTGTCCATATCATCATCTACCAGTGTCAACCTTACAGCAGTG  361

Qy   496  TTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGTCACCGCCACCACCAACA  555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362  TTCTCATCCTGGCCTTCATCAGCCTGGACCGGTACCTGTCACCGCCACCACCAACA  421

Qy   556  GTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGTCTGGATCCCAG  615
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422  GCCAGAGCGCGAGGAAGCTGCTGGCTGCGGAAAAGCCGTCTATGTGGGTGTCTGGATCCCG  481

Qy   616  CCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCATCAG  673
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   482  CCCTCCTCCTGACTATCCCTGACATCATCTTCGCCGACGTCAGCATGAGCCAGGGGACCGGCAG  539
```

```
RESULT 15
RNU90610
LOCUS        RNU90610               1050 bp    mRNA    linear   ROD 26-MAR-1997
DEFINITION   Rattus norvegicus CXC chemokine receptor (CXCR4) mRNA, complete
             cds.
ACCESSION    U90610
VERSION      U90610.1  GI:1906612
KEYWORDS     .
SOURCE       Rattus norvegicus (Norway rat)
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE    1  (bases 1 to 1050)
  AUTHORS    Harrison,J.K. and Salafranca,M.N.
  TITLE      Molecular cloning of rat CXCR4
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 1050)
  AUTHORS    Harrison,J.K. and Salafranca,M.N.
  TITLE      Direct Submission
  JOURNAL    Submitted (25-FEB-1997) Pharmacology and Therapeutics, University
             of Florida, 1600 SW Archer Rd., Gainesville, FL 32610-0267, USA
FEATURES             Location/Qualifiers
     source          1..1050
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /tissue_type="spleen"
     gene            1..1050
                     /gene="CXCR4"
     CDS             1..1050
                     /gene="CXCR4"
                     /codon_start=1
                     /product="CXC chemokine receptor"
                     /protein_id="AAB50408.1"
                     /db_xref="GI:1906613"
                     /translation="MEIYTSDNYSEEVGSGDYDSNKEPCFRDENENFNRIFLPTIYFI
                     IFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFG
                     KFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQSARKLLAEKAVVGVWIP
                     ALLTIPDIIFADVSQGDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVLSCYCI
                     IISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLEVIKQGCEFESV"
```

```
Qy   599  GGGGCTCTGAGGTCCCAGCCCCTCCTGCTATACCTGACTTCATCTTTGCCGACGTCAG  658
          ||||| ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db 21120  GGGTGTCTGGATCCCGCCCCTCCTCCTGACTATCCCTGACATCATCTTCGCCGATGTCAG  21179

Qy   659  CCAGGGGACCATCAG  673
          ||||||||| |||||
Db 21180  CCAGGGGGACCGGCAG  21194
```

```
RESULT 14
AF452185
LOCUS        AF452185               1050 bp    mRNA    linear   ROD 18-DEC-2001
DEFINITION   Rattus norvegicus strain Holtzman chemokine receptor CXCR4 (Cxcr4)
             gene, complete cds.
ACCESSION    AF452185
VERSION      AF452185.1  GI:17902280
KEYWORDS     .
SOURCE       Rattus norvegicus (Norway rat)
  ORGANISM   Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE    1  (bases 1 to 1050)
  AUTHORS    Simen,A.A. and Miller,R.J.
  TITLE      Chemokine regulation of neuronal signaling and gp120 neurotoxicity
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 1050)
  AUTHORS    Simen,A.A. and Miller,R.J.
  TITLE      Direct Submission
  JOURNAL    Submitted (25-NOV-2001) Dept. Psychiatry, Yale University, New
             Haven, CT 06508, USA
FEATURES             Location/Qualifiers
     source          1..1050
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Holtzman"
                     /db_xref="taxon:10116"
                     /sex="female"
                     /tissue_type="whole brain"
     gene            1..1050
                     /gene="Cxcr4"
     CDS             1..1050
                     /gene="Cxcr4"
                     /codon_start=1
                     /product="chemokine receptor CXCR4"
                     /protein_id="AAL47855.1"
                     /db_xref="GI:17902281"
                     /translation="MEIYTSDNYSEEVGSGDYDSNKEPCFRDENENFNRIFLPTIYFI
                     IFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFG
                     KFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPKKLLAEKAVVGVWIP
                     ALLTIPDIIFADVSQGDRYICDRLYPDSLWMVVFQFQHIMVGLILPGIVLSCYCI
                     IISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLEVIKQGCEFBSV
                     VHKWISITRALAFPHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKILSKGKRGGH
                     SSVSTESBSSSFHSS"
     variation       429..430
                     /gene="Cxcr4"
                     /note="compared to Wistar strain sequence presented in
                     GenBank Accession Number U90610"
                     /replace="cg"
ORIGIN
Query Match        69.0%;  Score 472.4;  DB 10;  Length 1050;
Best Local Similarity  92.4%;  Pred. No. 8.9e-114;
Matches  497;  Conservative  0;  Mismatches  41;  Indels  0;  Gaps  0;
```

```
Qy   136  TGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCA  195
          || ||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db     2  TGGAAATATACACTTCTGATAACTACTCTGAGAACTACTCCGAGAACTAGGTCTGAGACTATGACTCCA  61

Qy   196  ACAAGGAACCCTGCTTCCGGGATGAAAAGTCCATTTCAATAGGATCTTCCTGCCCACCA  255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62  ACAAGGAACCCTGCTTCCGGGATGAAAACGAAAAACTTCAACAGAATCTTCAACGGATCTTCCTGCCCACCA  121
```

ORIGIN

```
                VHKWISITEALAFFHCCLNPILYAFLGAKFKSAQHALNSMSRGSSLKILSKGKRGGH
                SSVSTESESSFHSS"

Query Match         68.5%;   Score 469.2;   DB 10;   Length 1050;
Best Local Similarity  92.0%;   Pred. No. 6.3e-113;
Matches 495;  Conservative   0;  Mismatches  43;  Indels   0;  Gaps   0;

Qy   136  TGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTCTGAGACTATGACTCCA  195
               ||||||| ||||||||| ||||||||||| |||||||| |||||||||||||||||||||
Db     2  TGGAAATATACACTTCCGATAACTACTCCGAAGAAGTAGGGTCTCTGAGACTATGACTCCA   61

Qy   196  ACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCA  255
               |||||||||||||||||||||  ||||||    ||  ||||| ||||||||||||||||
Db    62  ACAAGGAACCCTGCTTCCGGGATGAAAAACGAAAAACTTCAACAGGATCTTCCTGCCCACCA  121

Qy   256  TCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGG  315
               |||| ||| |||||||||||||||||||||| ||||||| | |||| |||||||||||||
Db   122  TCTATTTTATCATCTTCTTGACTGGCATAGTGGGCAATGGGTTGGTAATCCTGGTCATGG  181

Qy   316  GTTACCAGAAGAAGCTAAGGAGCATGACGGACAAGAAGTACCGGCTGCACCTGTCAGTGGCTG  375
               |||||||||||||||| |||||||||| ||||||||||||||||||||||||||| ||||||
Db   182  GTTACCAGAAGAAGCTGAGGAGCATGACAGACAAGAAGTACCGGCTGCACCTGTCCGTGGCTG  241

Qy   376  ACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACT  435
               |||||||||||||||||||||||||||||||||||||  || |||||||||||||||||||
Db   242  ACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTGGACGCCATGGCTGACTGGTACT  301

Qy   436  TTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCG  495
               |||||||||||| | ||||||||| |||||||||||||| ||||||||| |||||||| |
Db   302  TTGGGAAATTTTTATGTAAGGCTGTGCATATCATCTACACCCTCAACCTTTACAGCAGTG  361

Qy   496  TTCTCATCCTGGCCTTCATCAGCCTGACCCGGACCTCGCCATTGTCCAGCCACCAACA  555
               ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   362  TTCTCATCCTGGCCTTCATCAGCCTGACCCGGACCTTGCCATTGTCCAGCCACCAACA  421

Qy   556  GTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAG  615
               | ||||||  | ||||  |||||||||||||||||  ||||||||||||||||||||||
Db   422  GCCAGAGGCGCGGAGGAAGCTGCTGGCTGAAAAGGCCCGTCTATGTGGGCGTCTGGATCCCG  481

Qy   616  CCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGGACATCAG  673
               ||||||||||||||| |||||| ||||||| ||||| ||||  |||| ||||||| ||||
Db   482  CCCTCCTCCTGACTATCCCTGACATCATCTTCGCCGATGTCGCCGAGTGCAGCACGGGCAG  539
```

Search completed: May 17, 2004, 12:10:47
Job time : 2772.88 secs

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 03:46:58 ; Search time 293.749 Seconds
(without alignments)
9906.453 Million cell updates/sec

Title: US-09-367-052-5
Perfect score: 685
Sequence: 1 ccatcctaatacgactcact.........gacatcagtcaggggatga 685

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124098041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match	Length	DB	ID	Description
1	685	100.0	685	1877	2	AAV46370	Nucleic a
2	566	82.6		1758	2	AAZ27611	Mouse CXC
3	469.2	68.5		1050	9	ADB58421	Toxicity-
4	469.2	68.5		1050	9	ADB53005	Primary r
5	446.8	65.2		1679	4	AAD12801	Human neu
6	446.8	65.2		1679	4	AAI65467	Nucleotid
7	446.8	65.2		1679	6	ABS53992	DNA encod
8	446.8	65.2		1679	6	ABS53750	DNA encod
9	446.8	65.2		1679	6	ABV78039	Hypoxia-i
10	446.8	65.2		1679	7	ABX08779	Angiogene
11	446.8	65.2		1679	7	ABX74454	Human cDN
12	446.8	65.2		1679	7	ABZ68886	Nucleotid
13	446.8	65.2		1679	7	ABZ42642	Human CXC
14	446.8	65.2		1679	9	ADC98645	Human CXC
15	446.8	65.2		1679	6	ABZ35348	Human gen
16	446.8	65.2		1711	9	ADB47320	Human CDN
17	445.8	65.1		1664	9	AAA61656	Human CXC
18	445.8	65.1		1664	2	AAZ40014	CXCR4 cod
19	439.2	64.1		1737	2	AAQ29506	New plate
20	439.2	64.1		1737	2	AAQ80521	Human mon
21	439.2	64.1		1737	2	AAQ99007	Chemokine
22	437.6	63.9		1670	2	ABK83803	Human cDN
23	437.6	63.9		1670	6	ABN95645	Gene #214
24	437.6	63.9		1670	7	ACC46765	Human COP
25	432.4	63.1		1102	7	ACA64750	Chemokine
26	430.8	62.9		1059	5	AAZ38553	Human CXC
27	430.8	62.9		1059	5	ABI97982	Non-endog
28	430.8	62.9		1059	6	ABQ74938	Human CXC
29	430.8	62.9		1059	6	AAH99952	CXCR4 enc
30	430.8	62.9		1119	6	AAK98241	Human cys
31	430.8	62.9		1225	6	ABZ35630	Human gen
32	430.8	62.9		1225	6	ACA56637	Human sig
33	430.8	62.9		1588	2	AAZ27610	Human CXC
34	429.2	62.7		1059	7	ACC72707	Human can
35	428	62.5		1317	2	AAQ66179	Seven tra
36	428	62.5		1317	2	AAV18357	Human RM3
37	428	62.5		1317	6	AAA91726	Human 7TM
38	428	62.5		1944	2	ABK54255	Human 7 t
39	428	62.5		1944	2	AAX15882	cDNA enco
40	428	62.5		5161	3	AAA34774	Human ade
41	428	62.5		5161	3	AAF20896	Human CCR
42	428	62.5		5161	6	AAH99951	CXCR4 gen
43	428	62.5		5161	6	ABZ96590	Human CCR
44	428	62.5		5161	7	ACA64841	Human CXC
45	427.6	62.4		1582	6	ABS64793	CXCR4 rec

ALIGNMENTS

RESULT 1
AAV46370
ID   AAV46370 standard; cDNA to mRNA; 1877 BP.
XX
AC   AAV46370;
XX
DT   20-NOV-1998 (first entry)
XX
DE   Nucleic acid encoding a murine CXC chemokine receptor.
XX
KW   Mouse; CXC chemokine receptor; pre-B cell line DW34;
KW   CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1; HIV infection;
KW   screening; inhibitor; AIDS; ds.
XX
OS   Mus sp.
XX
FH   Key   Location/Qualifiers
FT   CDS   120..1199
FT   /*tag= a
XX
PN   WO9835035-A1.
XX
PD   13-AUG-1998.
XX
PF   07-FEB-1997; 97WO-JP000299.
XX
PR   07-FEB-1997; 97WO-JP000299.
XX
PA   (SHIO) SHIONOGI & CO LTD.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;
PI   Nakajima T, Yoshie O;
XX
DR   WPI; 1998-447232/38.
DR   P-PSDB; AAW64778.
XX
PT   Mouse CXC chemokine receptor binding to PBSF/SDF-1 pre-B cell stimulatory
PT   factor - is useful for screening of potential HIV infection and AIDS
PT   inhibitors.
XX
PS   Claim 3; Page 39-42; 76pp; Japanese.
XX
CC   The present sequence encodes a murine CXC chemokine receptor which binds
CC   to the mouse CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1. The
CC   nucleic acid is isolated from mouse pre-B cell line DW34. The receptor of
CC   and cells expressing it can be used in the study and mapping of the

```
CC    mechanism of HIV infection and in screening of potential inhibitors of
CC    HIV infection and the development of AIDS
XX
SQ    Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T; 0 U; 0 Other;

Query Match             100.0%;  Score 685;  DB 2;  Length 1877;
Best Local Similarity   100.0%;  Pred. No. 5.8e-191;
Matches  685;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CCATCCTAATACGACTCACTATAGGGCTCGAGCGCCGGGCCGGGCAGGTCGCAGGTAGCAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCATCCTAATACGACTCACTATAGGGCTCGAGCGCCGGGCCGGGCAGGTCGCAGGTAGCAG 60

Qy     61 TGACCCCTCGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTGTGAGGCGAGTGTTGCCA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGACCCCTCGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTGTGAGGCGAGTGTTGCCA 120

Qy    121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCCTGAAGAAGTGGGGTCTG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCCTGAAGAAGTGGGGTCTG 180

Qy    181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGA 240

Qy    241 TCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTGGGGCAATGGATTGG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTGGGGCAATGGATTGG 300

Qy    301 TGATCCTGGTCATGGGTACCAGAAGAAGCTAAGGAGCATGACGACGAAGTACCGGCTGC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGATCCTGGTCATGGGTACCAGAAGAAGCTAAGGAGCATGACGACGAAGTACCGGCTGC 360

Qy    361 ACCTGTCAGTGGCTGACCTCCTTCTTTGTCATCACACTCCCCTTCTCGGGCAGTGATGCCA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ACCTGTCAGTGGCTGACCTCCTTCTTTGTCATCACACTCCCCTTCTCGGGCAGTGATGCCA 420

Qy    421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCA 480

Qy    481 ACCTCTACAGACAGCGGTTCTCATCTGCCTTCATCAGCCTGGCTGAAAGGCAGTCTATGTGG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCTCTACAGACAGCGGTTCTCATCTGCCTTCATCAGCCTGGCTGAAAGGCAGTCTATGTGG 540

Qy    541 TCCACGCCACCAACAGTCAAGGCAAGGAAACTGCTGGCTGGCTGAAAGGCAGTCTATGTGG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCCACGCCACCAACAGTCAAGGCAAGGAAACTGCTGGCTGGCTGAAAGGCAGTCTATGTGG 600

Qy    601 GCGTCTGGATCCCAGCCCTCCTCCGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCGTCTGGATCCCAGCCCTCCTCCGACTATACCTGACTTCATCTTTGCCGACGTCAGCC 660

Qy    661 AGGGGGACATCAGTCAGGGGGATGA 685
            |||||||||||||||||||||||||
Db    661 AGGGGGACATCAGTCAGGGGGATGA 685


RESULT 2
AAZ27611
ID   AAZ27611 standard; DNA; 1758 BP.
XX
AC   AAZ27611;
XX
DT   16-DEC-1999  (first entry)
XX
DE   Mouse CXCR4 coding sequence.
XX
KW   CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;
KW   tissue repairing agent; vascularisation; ss.
XX
OS   Mus sp.
XX
```

```
PN   WO9948528-A1.
XX
PD   30-SEP-1999.
XX
PF   23-MAR-1999;  99WO-JP001448.
XX
PR   24-MAR-1998;  98JP-00095448.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
PA   (KISH/) KISHIMOTO T.
XX
PI   Kishimoto T,  Nagasawa T,  Tachibana K;
XX
DR   WPI; 1999-591042/50.
DR   P-PSDB; AAY39994.
XX
PT   CXCR4-potentiating agents and methods useful for inhibiting
PT   neovascularization, and treating solid cancers.
XX
PS   Disclosure; Page 50-51; 63pp; Japanese.
XX
CC   This sequence encodes the mouse CXCR4 protein. The invention relates to
CC   remedies inhibiting neovascularisation, remedies for solid cancer,
CC   remedies for diseases pathologically caused by neovascularisation and
CC   tissue repairing agents containing as the active ingredient a substance
CC   capable of potentiating CXCR4. Based on a finding that vascularisation is
CC   inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC   remedies inhibiting vascularisation which contain as the active
CC   ingredient a substance capable of potentiating CXCR4, remedies for solid
CC   cancer, remedies for diseases pathologically caused by neovascularisation
CC   and tissue repairing agents containing as the active ingredient a
CC   substance capable of potentiating CXCR4. It is also possible to establish
CC   methods for treatment with the use of these remedies
XX
SQ   Sequence 1758 BP; 447 A; 414 C; 369 G; 528 T; 0 U; 0 Other;

Query Match             82.6%;  Score 566;  DB 2;  Length 1758;
Best Local Similarity  100.0%;  Pred. No. 6.2e-156;
Matches  566;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    120 ATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCCTGAAGAAGTGGGGTCT 179
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCCTGAAGAAGTGGGGTCT 60

Qy    180 GGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGG 239
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGG 120

Qy    240 ATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTGGGCAATGGATTG 299
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTGGGCAATGGATTG 180

Qy    300 GTGATCCTGGTCATGGGTTACCAGAAGAGAGCTAAGGAGCATGACGACGAAGTACCGGCTG 359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGATCCTGGTCATGGGTTACCAGAAGAGAGCTAAGGAGCATGACGACGAAGTACCGGCTG 240

Qy    360 CACCTGTCAGTGGCTGACCTCCTTCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGATGCC 419
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CACCTGTCAGTGGCTGACCTCCTTCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGATGCC 300

Qy    420 ATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTC 479
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTC 360

Qy    480 AACCCTACAGCAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGCTGACCTCGCCATT 539
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AACCCTACAGCAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGCTGACCTCGCCATT 420

Qy    540 GTCCACGCCACCAACAGTCAAAGGCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTG 599
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTCCACGCCACCAACAGTCAAAGGCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTG 480

Qy    600 GGCGGTCGTCGATCCCAGCCCTCCTCCGACTATACCTGACTTCATCTTTGCCGACGTCAGC 659
```

```
Db     2 TGGAAATATACACTTCGGATAACTACTCCGAAGAAGTAGGGTCTCGAGACTATGACTCCA  61
         ||||||||||||||| ||||| ||||||||||||||||  ||||||||| ||||||||||
Qy   196 ACAAGGAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCCACCA  255
         |||||||||||||||||||||||||||   |  ||||| ||  |||||||||||||||||
Db    62 ACAAGGAACCCTGCTTCCGGATGAAAACGAAAACTTCAACAGGATCTTCCTGCCCCACCA  121

Qy   256 TCTACTTCATCATCTTCTTGACTAGTCGGCATAGTCGGCAATGGGTTGGTATCCTGGTCATGG  315
         |||| || |||||||||||||||| ||||||||| ||||||||||||||||||||||||||
Db   122 TCTATTTTATCATCTTCTTGACTGGCATAGTGGGCAATGGGTTGGTATCCTGGTCATGG  181

Qy   316 GTTACCAGAAGAAGCTAAGGAGCATGACGGACGAAGTACCGGCTGCACCGTGTCAGTGGCTG  375
         |||||||||||||||| ||||||||||||||| |||||||||||||||||||||| |||||||
Db   182 GTTACCAGAAGAAGCTGAGGAGCATGACAGACAGAAGTACCGGCTGCACCGTGTCCGTGGCTG  241

Qy   376 ACCTCCTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGACTGGTACT  435
         |||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   242 ACCTCCTTGTCATCACACTCCCCTTCTGGGCAGTGACGCCATGGCTGACTGGTACT  301

Qy   436 TTGGGAAATTTTGTGTAAGGCTGTCCATATCATCACTGTCAACCTCTACAGCAGCG  495
         ||||||||||||| ||||||||||||||||||||| ||||||||| ||||||||| 
Db   302 TTGGGAAATTTTTATGTAAGGCTGTCCATATCATCACCGTCAACCTTTACAGCAGTG  361

Qy   496 TTCTCATCCTGGCCTTCATCAGCCTGACCGGGACCTCGCCATTGTCCACGCCACCAACA  555
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   362 TTCTCATCCTGGCCTTCATCAGCCTGACCGGGACCTCGCCATTGTCCACGCCACCAACA  421

Qy   556 GTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGCAGTCTATGTGGGGTCTCGGATCCCAG  615
         | |||  || |||||| |||||||||||||||  |||||||||||||  ||| |||||| |
Db   422 GCCAGAGCGCGAGGAAGCTGCTGGCTGAAAAGGCCGTCTATGTGGGTCTCTGGATCCCCG  481

Qy   616 CCCTCCTTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCAGCGGGACCATCAG  673
         ||||||||||||||||| ||||| |||||||| |||||  |||||| ||||||| |||||
Db   482 CCCTCCTTCCTGACTATCCCTGACATCATCTTCTGCCGGATGTCAGCCAGCGGGACGGGCAG  539
```

RESULT 4

```
ID   ADB53005
ID   ADB53005 standard; DNA; 1050 BP.
XX
AC   ADB53005;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3547.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
PR   19-APR-2002; 2002US-0373601P.
PR   19-APR-2002; 2002US-0373602P.
PR   22-APR-2002; 2002US-0374139P.
PR   08-MAY-2002; 2002US-0378370P.
PR   09-MAY-2002; 2002US-0378652P.
PR   09-MAY-2002; 2002US-0378653P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-JUL-2002; 2002US-0394230P.
PR   09-JUL-2002; 2002US-0394253P.
```

```
Db   481 GGCGTCTGGATCCCAGCCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGC  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   660 CAGGGGCACATCAGTCAGGGGATGA  685
         |||||| |||||||||||||| ||||
Db   541 CAGGGGGACATCAGTCAGGGGGATGA  566
```

RESULT 3

```
ID   ADB58421
ID   ADB58421 standard; DNA; 1050 BP.
XX
AC   ADB58421;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Toxicity-related gene, SEQ ID 3447.
XX
KW   Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW   drug screening; toxicity assay; ds.
XX
OS   Unidentified.
XX
PN   WO2003064624-A2.
XX
PD   07-AUG-2003.
XX
PF   31-JAN-2003; 2003WO-US003194.
XX
PR   31-JAN-2002; 2002US-00060087.
PR   15-MAR-2002; 2002US-0364045P.
PR   15-MAR-2002; 2002US-0364055P.
PR   30-DEC-2002; 2002US-0436643P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
DR   WPI; 2003-689530/65.
XX
PT   Predicting a toxic effect of a compound, useful in identifying toxicity
PT   markers in liver tissues or cells for drug screening and toxicity assays,
PT   comprises preparing gene expression profile of tissue or cells exposed to
PT   the compound.
XX
PS   Claim 1; SEQ ID NO 3447; 1156pp; English.
XX
CC   The present invention relates to a method for predicting a toxic effect
CC   of a compound. The method comprises preparing a gene expression profile
CC   of a tissue or cell sample exposed to the compound, and comparing the
CC   gene expression profile to a database comprising SEQ ID 1-4925, where
CC   differential expression of the gene indicates at least one toxic effect.
CC   The method is useful for predicting at least one toxic effect of a
CC   compound, predicting hepatotoxicity or the progression of a toxic effect
CC   of a compound, identifying an agent that modulates the onset or
CC   progression of a toxic response, predicting the cellular pathways that a
CC   compound modulates in a cell, and identifying an agent that modulates at
CC   least one activity of a protein. The method and compositions of the
CC   present invention using a database of genes having liver toxin-induced
CC   differential expression, are useful in identifying toxicity markers in
CC   liver tissues or cells for drug screening and toxicity assays. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1050 BP; 231 A; 313 C; 247 G; 259 T; 0 U; 0 Other;

Query Match             68.5%;  Score 469.2;  DB 9;  Length 1050;
Best Local Similarity   92.0%;
Matches  495;  Conservative   0;  Mismatches  43;  Indels   0;  Gaps   0;

Qy   136 TGAGTATATACACTTCGATAACTACTCTGATAACTACTCTGAAGAAGTGGGGTCTCGGAGACTATGACTCCA  195
         ||  |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

PR   04-SEP-2002; 2002US-0407688P.
PR   28-JAN-2003; 2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3547; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 1050 BP; 231 A; 313 C; 247 G; 259 T; 0 U; 0 Other;

Query Match   68.5%; Score 469.2; DB 9; Length 1050;
Best Local Similarity   92.0%; Pred. No. 1.6e-127;
Matches 495; Conservative   0; Mismatches 43; Indels   0; Gaps   0;

Qy   136 TGAGTAATATACACTTCGATAATAACTACTCTGAAGAAGTGGGGTCTGAGAGACTATGACTCCA 195
              | ||||||||||||||||| ||||||| || ||||||||||||||||| |||||||||||||
Db   2 TTGAAATATACACTTCGATAACTACTCTCCGAAGAAGTAGGGTCTGAGAGACTATGACTCCA 61

Qy   196 ACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCA 255
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   62 ACAAGGAACCCTGCTTCCGGGATGAAAACGAAAACTCAACAGGATCTTCCTGCCCACCA 121

Qy   256 TCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTCATCCTGGTCATGG 315
              |||| | ||||||||||||||||||||| ||||||||||||||| |||||| |||||||||
Db   122 TCTATTTTATCATCTTCTTGACTGGCATAGTGGGCAATGGTTGGTAATCCTGGTCATGG 181

Qy   316 GTTACCAGAAGAGCTAAGGAGCTATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTG 375
              |||||||||||||| |||||||| ||||||||||||||||||||| |||||||||||||
Db   182 GTTACCAGAAGAGCTGAGGAGCTGCAGACAGAAGTACCGGCTGCACCTGTCCGTGGCTG 241

Qy   376 ACCTCCTCTTTGTCATCACACTCCCCTCTGGCAGTTGATGCCATGGCTGACTGGTACT 435
              ||||||||||| ||||||||||||||||||||||||| ||||||| |||||||||||||
Db   242 ACCTCCTCTTTGTCATCACACTCCCCTCTCTGGCAGTGGACGCCATGGCTGACTGGTACT 301

Qy   436 TTGGGAAATTTTTGTGTAAGGCTGCCATATCATCATCTACACTGTCAACCTCTACAGCAGCG 495
              ||||||||||||| |||||||||||| |||||||||||||||||||||| ||| |||||||
Db   302 TTGGGAAATTTTTATGTAAGGCTGTGTCATATCATCATCTACACCGTCAACCTCTTTACAGCAGTG 361

Qy   496 TTCTCATCCTGGCCTTCATCAGCCTGGACCCGGTACCTGCCCATTGTCCACGCCACCAACA 555
              ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   362 TTCTCATCCTGGCCTTCATCAGCCTGGACCCGGTACCTGCCCATTGTCCACGCCACCAACA 421

Qy   556 GTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTTGGATCCCAG 615
              ||||||||||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db   422 GCCAGACGGCCGGAGGAAGCTGCTGGCTGCGTCTATGTGGGTGTCTTCGGATCCCCG 481

Qy   616 CCCTCCTCTGACTATACCTGACTTCATCTTTGCCGACGTGCAGCCAGGGGACATCAG 673
              ||||||||| ||||| ||||||||||||||||| |||| ||||| |||||||||||||||
Db   482 CCCTCCTCCTGACATCCCTGACATCATCTTCCGCCGATGTGCAGCCAGGGGACGGCAG 539

```
CC   The specification describes an antibody which is specific for an
CC   antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or
CC   is exposed upon gp120 binding of DC-SIGN due to concomitant
CC   conformational change. DC-SIGN is a receptor that is specifically
CC   expressed on dendritic cells and facilitates infection of T lymphocytes
CC   with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC
CC   -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)
CC   with high affinity. The antibody of the invention inhibits the trans
CC   enhancement of HIV entry into a T cell or macrophage facilitated by
CC   dendritic cells. The antibody is useful to treat or prevent HIV
CC   infection. The present sequence represents a human polynucleotide, which
CC   is used in the course of the invention
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
```

```
Query Match             65.2%;  Score 446.8;  DB 4;  Length 1679;
Best Local Similarity   83.0%;  Pred. No. 7.7e-121;
Matches 523;  Conservative   0;  Mismatches 102;  Indels   5;  Gaps   1;

Qy         44 GCAGGTGCAGGTAGCAGTGACCCTCGAGGCGTTGTGCTCCGGTAACCACCACGGCTG 103
              ||| ||||||||||||| |||||| ||| || |||||||||||||| ||||| |||
Db         12 GCGGCAGCAGGTAGCAAGTGACGACGGGGCCTCGAGTGCTCCAGTGCCACCGCATCTG 71

Qy        104 TAGAGCAAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
              |||||||||||| |||||||||   | |||    |||||||||||||| |||||||||||
Db         72 GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy        164 TGAAGAAGTGGGGTCTGAGACATGACTCAACAACAAGGAACCCTGCTTCCGGGATGAAAA 223
              ||||||| ||||| |||||| ||||| || || |||||||||||| ||||| |||||||
Db        127 CGAGGAAATGGGCTCAGGGGACTATGACACCATGAAGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy        224 CGTCCATTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
              ||| ||||||||| | |||||||||| ||||||||||| ||||||||||||| |||||||
Db        187 TGCTAATTCAATAAAATCTTCCTGCCCACCATCTACTACTCCATCATCTTCTTAACTGGCAT 246

Qy        284 AGTCGGCAATGAATGGTGATCCTGGTCATGGTCATGGGCAATGAAGCTAAGGAGCATGAC 343
              ||| |||||||||||| | |||||||||||||||||||||||||||  |||| |||||||
Db        247 TGTGGGCAATGAATGGTTGTCATCCTGGTCATGGTCATGGGCAATGAAAACTGAGAAGCATGAC 306

Qy        344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTTGTCATCACACTCCCCTT 403
              |||||||||| ||||||||||||||||| |     |||||||||||||| |||||||
Db        307 GGACAAGTACCAGGCTGCACCTGTCAGTGCCAGGCGCTCCTTCTTGTCATCACGCTTCCCTT 366

Qy        404 CTGGCAGTTGATGCCATGGCCGTTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGA 463
              |||| ||||||||||| ||| | |||| | |||||||| ||||| |||||||||||||
Db        367 CTGGGCAGTTGATGCCGTGCGCAAACTGGTACTTTGGGACTTGCAAGGCAGTCCA 426

Qy        464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGA 523
              ||||||||||||| |||||||||| |||||| ||| |||||||||||||||||||||
Db        427 TGTCATCTCACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGCCTGA 486

Qy        524 CCGGTACCTGCCATTGTCCACGCCACCAACAGTCAGAGGCCTGGTCCGGCTGA 583
              ||||||| ||||||| |||| || |||||||| |||||||||| ||| |||||
Db        487 CCGCTACCTGGCCATCGTCCACCGCCACCAACAGTCAGAGGCCAGGGAGCTGTGGCTGA 546

Qy        584 AAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGCTATACCTGACTTCAT 643
              |||| ||||| ||||||| |||||||||| ||| |||||||||| ||||| ||||||
Db        547 AAAGGTGGTTCTATGTTGGCGTCTGGATCCCTGGTCGGATCCCTGGCCTATTCCCGACTTCAT 606

Qy        644 CTTTGCCGACGTCAGCCAGGGGGACATCAG 673
              |||||||||||| ||||||||| || |||
Db        607 CTTTGCCAACGTCAGTGAGGCAGATGACAG 636
```

```
RESULT 6
AAI65467
ID   AAI65467 standard; DNA; 1679 BP.
XX
AC   AAI65467;
XX
DT   10-DEC-2001  (first entry)
DE   Nucleotide sequence of a human polynucleotide.
XX
KW   Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;
KW   C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.
XX
OS   Homo sapiens.
XX
PN   WO200164752-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US006322.
XX
PR   02-MAR-2000; 2000US-00517605..
XX
PA   (UYNY ) UNIV NEW YORK STATE.
PA   (UYNI-) UNIV NIJMEGEN.
XX
PI   Littman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;
XX
DR   WPI; 2001-602565/68.
XX
PT   An antibody for the treatment or prevention of HIV-infection comprises a
PT   gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of
PT   DC-SIGN due to concomitant conformational change.
XX
PS   Disclosure; Page 126-127; 131pp; English.
XX
```

```
RESULT 7
ABS53992
ID   ABS53992 standard; cDNA; 1679 BP.
XX
AC   ABS53992;
XX
DT   02-DEC-2002  (first entry)
XX
DE   DNA encoding human CXC chemokine receptor 4 (CXCR4).
```

```
Db     12 GGGCAGCAGGTAGCAAAGTGAGCCGAGGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG   71
          ||| ||||||||||||| |||||| ||||   |||||| |||||||||| ||| |||||||

Qy    104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC  163
          |||||| ||||||| |||||||||||    ||   |||||||||||||||| ||||||||
Db     72 GAGAACCAGCGGTTACCATGGA----GGGGATCAGTATATACACTTCAGATAACTACAC  126

Qy    164 TGAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGAACCCTGCTTCCGGGATGAAAA   223
          |||||| ||||| |||||| |||||||| |||||||||||||  |||  |  |||||||
Db    127 CGAGGAAATGGGCTCAGGGACTATGACTCCATGGAAGAACCCTGTTTCCGTGAAGAAAA  186

Qy    224 CGTCCATTCAAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCAT  283
          |||||  |||||   |||||||||||||||||||||||||||||||||| || ||||||
Db    187 TGCTAATTTCAATAAATCTTCGCCCACCATCTACTTCATCATCTTCTTAACTGGCAT   246

Qy    284 AGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC  343
          ||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db    247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAAACTGAGAGCATGAC   306

Qy    344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCTT  403
          ||||||||| ||||||||||||||||||||||||||||| ||||||||| ||||||||
Db    307 GGACAAGTACAGGCTGCACCTGTCAGTGCCGACCTCCTCTTTGTCATCACGCTTCCCTT  366

Qy    404 CTGGGCAGTTGATGCCAGTGCCGACTACTTTGGAAATTTTGTGTAAGGCTGTCCA  463
          ||||||||| ||||||| ||||| |||||||| |  ||  ||| |||| |||||||
Db    367 CTGGGCAGTTGATGCCGTGCCAAACTGGTACTTTGGGACTTCCTATGCAAGGCAGTCCA  426

Qy    464 TATCATCTCACACTGTCAACCTCCTCACAGCAGGTGTTCTCATCCTGGCCTTCATCAGCCTGA  523
          ||| || || || |||| ||||| |||||||||| || ||||||||||||||||||||| ||||
Db    427 TGTCATCTCACACAGTCAACCTCACAGCAGTGTTCATCCTGGCCTTCATCAGTCTGA  486

Qy    524 CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA  583
          |||| |||| |||||| |||||||| |||||||||| ||||||||||||| |||||||||
Db    487 CCGCTACCTCGCCATGCTCCACGCCACCGCCAACAGTCAGAGGCCAAGGAAGCTGTTGGGCTGA  546

Qy    584 AAAGGCAGTCTATGTGGGCGTCTATGTGGGCGTCCAGCCCTCCTGACTATACCTGACTTCAT  643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    547 AAAGGTGGTCTATGTTGGCGTCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCAT  606

Qy    644 CTTTGCCGACGTCAGCGCCAGGGGGACATCAG  673
          |||| ||||  |||||||||||  ||| ||||
Db    607 CTTTGCCAACGTCAGTGAGGCAGCAGATGACAG  636
```

XX
KW   CXC chemokine receptor 4; CXCR4; chemotaxis induction; SDF1a;
KW   stromal cell-derived factor 1alpha; angiogenesis; atherosclerosis;
KW   restenosis; ischaemic stroke; spinal cord injury; infection; ulcer;
KW   human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;
KW   AIDS; pain; cancer; benign prostatic hypertrophy; diabetes; obesity;
KW   anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW   hypotension; hypertension; urinary retention; osteoporosis; stroke;
KW   anginal pectoris; myocardial infarction; benign prostatic hypertrophy;
KW   migraine; vomiting; psychotic disorder; neurological disorder; anxiety;
KW   schizophrenia; dyskinesia; Huntingdon's disease; restenosis;
KW   inflammatory disease; rheumatoid arthritis; diabetic retinopathy;
KW   inflammatory bowel disease; atherosclerosis; Alzheimer's disease;
KW   congestive heart failure; cardiac remodeling; angiogenic diseases;
KW   solid tumour; Kaposi Sarcoma; human; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   89..1147
FT   /*tag= a
FT   /product= "CXCR4"
FT   /note= "CXC chemokine receptor 4"
XX
PN   US2002107195-A1.
XX
PN   2001US-00953692.
PD   08-AUG-2002.
PF   17-SEP-2001; 2001US-00953692.
XX
PR   21-JUL-1998;   98US-0093596P.
PR   21-JUL-1999;   99US-00358624.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI   Gupta SK;
XX
DR   WPI; 2002-697879/75.
DR   P-PSDB; ABG32977.
XX
PT   Inducing chemotaxis of endothelial cells, useful for regulating
PT   angiogenesis, e.g. for treating cancer, comprises treatment with stromal
PT   cell-derived factor 1 alpha.
XX
PS   Disclosure; Fig 1; 26pp; English.
XX
CC   The invention describes a method of inducing chemotaxis of endothelial
CC   cells by treatment with stromal cell-derived factor 1alpha (SDF1a). The
CC   method is used for stimulating EC chemotaxis, and thus angiogenesis, and
CC   is used for treating atherosclerosis, restenosis, ischaemic stroke and
CC   spinal cord injury. Inhibition of this process is useful in treatment and
CC   prevention of a very wide range of diseases, such as, infection (by
CC   bacteria, fungi, protozoa or viruses such as human immunodeficiency virus
CC   (HIV) and acquired immunodeficiency syndrome (AIDS)), pain, cancer and
CC   benign prostatic hypertrophy, diabetes, obesity, anorexia, bulimia,
CC   asthma, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, anginal pectoris,
CC   myocardial infarction, stroke, ulcers, benign prostatic hypertrophy,
CC   migraine, vomiting, psychotic and neurological disorders (e.g. anxiety,
CC   schizophrenia) and dyskinesias (e.g. Huntingdon's disease), inflammatory
CC   diseases, rheumatoid arthritis, diabetic retinopathy, inflammatory bowel
CC   disease, atherosclerosis, restenosis, Alzheimer's disease, congestive
CC   heart failure, cardiac remodeling, angiogenic diseases, solid tumours,
CC   and Kaposi Sarcoma. This sequence encodes the human CXC chemokine
CC   receptor 4 (CXCR4) polynucleotide.
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   65.2%;   Score 446.8;   DB 6;   Length 1679;
Best Local Similarity   83.0%;   Pred. No. 7.7e-121;
Matches 523;   Conservative   0;   Mismatches 102;   Indels   5;   Gaps   1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTG 103

```
XX
PN   US2002107196-A1.
XX
PD   08-AUG-2002.
XX
PF   17-SEP-2001; 2001US-00953717.
XX
PR   21-JUL-1998;   98US-00935956P.
PR   21-JUL-1999;   99US-0035862A.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Gupta SK;
XX
DR   WPI; 2002-706230/76.
DR   P-PSDB; ABG33065.
XX
PT   Inducing chemotaxis of endothelial cells, useful for regulating
PT   angiogenesis, e.g. for treating cancer, comprises treatment with stromal
PT   cell-derived factor 1 alpha.
XX
PS   Disclosure; Fig 1; 26pp; English.
XX
CC   The present invention relates to a new method for inducing chemotaxis of
CC   endothelial cells (EC). The method of the invention involves treatment
CC   with a stromal cell-derived factor 1a (SDF1alpha). The method is used for
CC   stimulating EC chemotaxis, and thus angiogenesis, and is used for
CC   treating atherosclerosis, restenosis, ischaemic stroke and spinal cord
CC   injury, while inhibition of this process is useful in treatment and
CC   prevention of a very wide range of diseases, e.g. infection (by bacteria,
CC   fungi, protozoa or viruses), pain, cancer and benign prostatic
CC   hypertrophy, diabetes, obesity, anorexia, bulimia, asthma, allergies,
CC   Parkinson's disease, acute heart failure, hypotension, hypertension,
CC   urinary retention, osteoporosis, angina pectoris, myocardial infarction,
CC   stroke, ulcers, migraine, vomiting, psychotic and neurological disorders
CC   and dyskinesias, inflammatory disorders, rheumatoid arthritis, diabetic
CC   retinopathy, inflammatory bowel disease, atherosclerosis, restenosis,
CC   Alzheimer's disease, congestive heart failure, cardiac remodelling,
CC   angiogenic diseases, solid tumours, and Kaposi sarcoma. The present
CC   nucleic acid sequence encodes the human CXCR4 protein of the invention.
CC   Note: This sequence represents SEQ ID NO 1 given in figure 1 of the
CC   specification
XX
SQ   Sequence 1679 BP; 407 A; 399 C; 367 G; 504 T; 0 U; 2 Other;

Query Match             65.2%;  Score 446.8;  DB 6;  Length 1679;
Best Local Similarity   83.0%;  Pred. No. 7.7e-121;
Matches  523;  Conservative  0;  Mismatches  102;  Indels  5;  Gaps  1;

Qy     44  GCAGTGCAGGTAGCAGTGACCCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG   103
Db     12  GCGGCAGCAGCAGTAGCAAAGTGACGCCGCAGGGGCCTGAGTGCTCCAGTAGCACCGCATCTG   71
Qy    104  TAGAGCGAGTGTTGCCATCGAACCGATCATGTGTGAGTATATACACTTCTGATAACTACTC   163
Db     72  GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTAGTATATACACTTCAGATAACTACAC   126
Qy    164  TGAAGAAGTGGGGTCTGGAGACTATGACTTCCAACAAGGAACCCTGCTTCCGGGATGAAAA   223
Db    127  CGAGGAAAATGGGCTCAGGGGACTATGACTTCATCGAAGGAACCCCTGTTTCCGTGAAGAAAA   186
Qy    224  CGTCCATTCAATAGGATCTTCCTGCCCACCACTACTTCATCATCTTCTTGACTGGCAT   283
Db    187  TGCTAATTCAATAAAATCTTCCTGCCCACCATCTACTTCTTCATCATCTCTTAACTGGCAT   246
Qy    284  AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGTAAGGAGCATGAC   343
Db    247  TGTGGGCAATGGATTGGTCATCCCGGTCATGGGTTACCAGAAGAAGAAACTGAGAAGCATGAC   306
Qy    344  GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCCTTTTGTCATCACACTCCCCTT   403
Db    307  GGACAAGTACAGGCTGCACCTGTGGCCGACCTCCCTTTTGTCATCACGGCTTCCCTT   366

Qy    404  CTGGGCAGTGATGCATGGCTGACGCGGTACTTGGGAAATTTTGTGTAAGGCTGTCCA   463
Db    367  CTGGGCAGTTGATGCGTGGCAAACTGGTACTTTGGAACTTCCATGCAAGGCAGTCCA   426
Qy    464  TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTCGGCCTTCATCAGCCTGGA   523
Db    427  TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTCGGCCTTCATCAGTCGGA   486
Qy    524  CCGGTACCTCGGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA   583
Db    487  CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGTGTTGGCTGA   546
Qy    584  AAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCCTCCTGACTATACCTGACTTCAT   643
Db    547  AAAGGTGGTCTATGTTGGCGTCTGGATCCCTGCCCCTCCTGCTGACTATTCCCGACTTCAT   606
Qy    644  CTTTGCCGACTTCAGCCAGGGGGACATCAG   673
Db    607  CTTTGCCAACGTCAGTGAGGCAGATGACAG   636

RESULT 9
ABV78039
ID   ABV78039 standard; DNA; 1679 BP.
XX
AC   ABV78039;
XX
DT   12-NOV-2002  (first entry)
XX
DE   Hypoxia-regulated protein coding sequence #59.
XX
KW   Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
KW   antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;
KW   hypoxia; tumourigenesis; angiogenesis; apoptosis; cancer;
KW   ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
KW   preeclapmsia; atherosclerosis; inflammatory condition; wound healing;
KW   inflammation; erythropoiesis; hair loss; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200246465-A2.
XX
PD   13-JUN-2002.
XX
PF   10-DEC-2001; 2001WO-GB005458.
XX
PR   08-DEC-2000; 2000GB-00030076.
PR   08-FEB-2001; 2001GB-00003156.
PR   25-OCT-2001; 2001GB-00025666.
XX
PA   (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI   White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
PI   Rayner WN;
XX
DR   WPI; 2002-627238/67.
XX
PT   Identifying a gene involved in disease for treating hypoxia-regulated
PT   conditions, comprises comparing the transcriptome/proteome of two cell
PT   types under different conditions and identifying a differentially
PT   regulated gene.
XX
PS   Claim 37; Page 378; 538pp; English.
XX
CC   The present invention relates to methods for identifying genes and
CC   proteins that are implicated in a specific disease or physiological
CC   condition. The method comprises comparing the transcriptome/proteome of a
CC   specialised cell type implicated in a disease or condition with that of a
CC   second specialised cell type, under two experimental conditions, and
CC   identifying a gene that is differentially regulated in the two
CC   specialised cell types under experimental conditions. ABV77873-ABV78116
CC   and ABP65061-ABP65257 were identified using the methods of the invention.
CC   The coding sequences and proteins are useful for treating a disease in a
```

```
XX
OS   Homo sapiens.
XX
PN   WO20279492-A2.
XX
PD   10-OCT-2002.
XX
PF   14-FEB-2002; 2002WO-US004915.
XX
PR   14-FEB-2001; 2001US-00784356.
PR   22-FEB-2001; 2001US-00791390.
PR   19-APR-2001; 2001US-0285475P.
PR   03-AUG-2001; 2001US-0310025P.
PR   13-NOV-2001; 2001US-0350666P.
PR   29-NOV-2001; 2001US-0334244P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Murray R, Glynne R, Watson SR, Aziz N;
XX
DR   WPI; 2003-040681/03.
DR   P-PSDB; ABU03495.
XX
PT   Detecting angiogenesis-associated transcript in a cell for diagnosing and
PT   treating cancer by contacting a sample with a polynucleotide that
PT   exhibits changes in expression level as a function of time in tissue
PT   undergoing angiogenesis.
XX
PS   Example 2; Page 217-218; 291pp; English.
XX
CC   The present invention relates to methods and compositions for detecting
CC   an angiogenesis-associated transcript in a cell in a patient. The method
CC   involves contacting a biological sample from the patient with a
CC   polynucleotide that selectively hybridises to a sequence at least 80%
CC   identical to any of the angiogenesis-associated human polynucleotide
CC   sequences given in the specification. These angiogenesis-associated
CC   polynucleotide sequences comprise genes that exhibit changes in
CC   expression levels as a function of time in tissue undergoing
CC   angiogenesis. The method and the polynucleotide sequences of the
CC   invention are useful for diagnosing and treating angiogenesis and
CC   angiogenesis-associated diseases e.g. cancer. The polynucleotide
CC   sequences are also useful in the gene therapy of such disorders. The
CC   angiogenesis-associated proteins encoded by the polynucleotide sequences
CC   are useful as a vaccine for therapeutic and prophylactic immunisation.
CC   ABX08739-ABX08853 represent angiogenesis-associated polynucleotide
CC   sequences
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match             65.2%;  Score 446.8;  DB 7;  Length 1679;
Best Local Similarity   83.0%;  Pred. No. 7.7e-121;
Matches  523;  Conservative  0;  Mismatches  102;  Indels  5;  Gaps  1;

Qy     44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
          ||| || |||||||||||| |||  |  ||| ||| ||||| || || ||  || |||
Db     12 GCGGCAGCAGGTAGCAGTGAGCCGCAGGGCCTGAGTGCTGCTCCAGTAGCCACCGCATCTG 71

Qy    104 TAGAGGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
          | |||  ||| ||||||||||||     | ||||  || |||||||||||| |||||||| |
Db     72 GAGAACCAGCCGGTTACCATGGA-----GGGGATCAGTAGTATATACACTTCAGATAACTACAC 126

Qy    164 TGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
          ||||| |||| ||||| || |||||||||||| ||||||||| |||| ||| ||| ||
Db    127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTCCGTGAAGAAAA 186

Qy    224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
          ||  ||| |||||  ||||||| ||||||||||||| |||||||||||||| |||||||
Db    187 TGCTAATTTCAATAAAAATCTTCCGCCCACCATCTACTCCATCATCTTCTTAACTGGCAT 246

Qy    284 AGTCGGCAATGGATTGGTGATCCTGGTACTTGGTACTTTGGTAAGGAGGCATGAC 343
          | | |||||||||||||| |||||||| |  |||| |  |||| |||||| ||||
Db    247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTCATGGGTTACCGAAGCAGTGAC 306
```

```
CC   patient, for manufacture of a medicament for treating hypoxia-regulated
CC   conditions, and for regulating tumourigenesis, angiogenesis, apoptosis,
CC   biological response to hypoxia conditions, or hypoxic-associated
CC   pathology in a patient. The coding sequences and proteins are also useful
CC   for monitoring the therapeutic treatment of a disease or physiological
CC   condition, such as cancer, ischaemic conditions, reperfusion injury,
CC   retinopathy, neonatal stress, preeclapmsia, atherosclerosis, inflammatory
CC   conditions, wound healing, inflammation, erythropoiesis or hair loss
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match             65.2%;  Score 446.8;  DB 6;  Length 1679;
Best Local Similarity   83.0%;  Pred. No. 7.7e-121;
Matches  523;  Conservative  0;  Mismatches  102;  Indels  5;  Gaps  1;

Qy     44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
          ||| || |||||||||||| |||  |  ||| ||| ||||| || || ||  || |||
Db     12 GCGGCAGCAGGTAGCAGTGAGCCGCAGGGCCTGAGTGCTGCTCCAGTAGCCACCGCATCTG 71

Qy    104 TAGAGGGAGTGTTGCCATGGAACCGATCAGTGTGAGTGAGTATATACACTTCTGATAACTACTC 163
          | ||||  ||| ||||||||||||     | ||||  ||| |||||||||||| |||||||| |
Db     72 GAGAACCAGCCGGTTACCATGGA------GGGGATCAGTAGTATATACACTTCAGATAACTACAC 126

Qy    164 TGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
          ||||| |||| ||||| || |||||||||||| ||||||||| |||| ||| ||| ||
Db    127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTCCGTGAAGAAAA 186

Qy    224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTAACTGGCAT 283
          ||  ||| |||||  ||||||| ||||||||||||| |||||||||||||| |||||||
Db    187 TGCTAATTTCAATAAAAATCTTCCGCCCACCATCTACTCCATCATCTTCTTAACTGGCAT 246

Qy    284 AGTCGGCAATGGATTGGTGATCCTGGTACTTGGTACTTTGGTAAGGCTGTGTCCA 343
          | | |||||||||||||| |||||||| |  |||| |  |||| ||| || ||||
Db    247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTCATGGGTTACCGAAGCAGTCCA 306

Qy    344 GGACAAGTACCCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTT 403
          |||||||||   ||||||||||||||||||| ||||||||||||||||||| ||||| ||
Db    307 GGACAAGTACAGAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCTCCCTT 366

Qy    404 CTGGCCAGTTGATGCCATGGCCGTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
          |||| |||||||||| |||||||| ||  ||||||||||||| ||  |||| |||||||||
Db    367 CTGGGCAGTTGATGCCGTGGCCGTCAAACTGTACTTTGGGAACTTCCATGCAAGGCAGTCCA 426

Qy    464 TATCATCACTGTCAACCTCTACAGCAGCCTTCTCATCCTGGCCTTCATCAGCCTGGA 523
          | |||||||| ||||||||||||||||  |||||||||||||||||||||| ||||
Db    427 TGTCATCTACACAGTCAACCTCTACAGCAGTCGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy    524 CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
          ||| ||||| ||||| |||||||||||||||||||| ||||||||||||| ||| |||||
Db    487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAAGGGCCAAGGAAACCCTGTTGGCTGA 546

Qy    584 AAAGGCAGTCTATGTGGGCGTCTGGATCCCGTCTGGATCCCTCCTGCTGGTATACCTGACTTCAT 643
          |||||  ||||||||  ||||||||||||| |||||||||||| ||||| || | ||||||||
Db    547 AAAGGTGGTCTATGTTGGCGTCTGGATCCCTGCCCTGCTGACTATTCCCGACTTCAT 606

Qy    644 CTTTGCCGAGTCGAGCCAGGGGGACATCAG 673
          ||||||| ||||  || || ||| ||| ||
Db    607 CTTTGCCAAGTCAGTCAGTGAGGCAGATGACAG 636
```

```
RESULT 10
ABX08779
ID   ABX08779 standard; cDNA; 1679 BP.
XX
AC   ABX08779;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Angiogenesis-associated human polynucleotide sequence #41.
XX
KW   Human; angiogenesis-associated transcript; angiogenesis;
KW   angiogenesis-associated disease; cancer; cytostatic; gene therapy; gene;
KW   ss.
```

CC   expression profiling of tumour and normal tissues. The present sequence
CC   represents a human cDNA sequence up-regulated in CC-RCC patients
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   65.2%; Score 446.8; DB 7; Length 1679;
Best Local Similarity   83.0%; Pred. No. 7.7e-121;
Matches 523; Conservative   0; Mismatches 102; Indels   5; Gaps   1;

```
Qy     44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGTCTCCGGTAACCACCACGGCTG 103
          |||||||||||||||       ||||   |||||||  ||||| |||    |  ||  ||
Db     12 GCGGCAGCAGGTAGCAAAGTGACGTGACGGCGGAGGGGCTGAGTGCTGCCACCGCCATCTG 71

Qy    104 TAGAGCCAGTGTTGCCATGGAACCGATCAGTGAGTATATACACTCTGATAACTACTC 163
          ||||| |||   ||| ||||                ||||| |||| ||||| |||||
Db     72 GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTATGATACACTTCAGATAACTACAC 126

Qy    164 TGAAGAAGTGGGGTCTGGAGACTATGATCCAACAAGGAACCTGCTTCCGGGATGAAAA 223
          |||| ||||||| ||  |||||||||| ||| ||||| ||      |||||| ||||||
Db    127 CGAGGAAATGGGCTCAGGGACTATGAGTCCATGAAGGGACTAGAAGAAACCCGTGTTCCGTGAAGAAAA 186

Qy    224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTTCATCATCCTTCTTGACTGGCAT 283
          ||| ||||||||   |||||||||||||||||  ||||||||| ||||| ||||||||
Db    187 TGCTAATTCAATAAAATCTTCCTGCCCACCATCACTTCCATCATCTTCTTTAACTGGCAT 246

Qy    284 AGTCGGCAATGGATTGGTGATCCTGGTTCATGGGTTACCAGGAGAAGGAGCTAAGGAGCATGAC 343
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGAC 306

Qy    344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTT 403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 GGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCATCACGTTCCTT 366

Qy    404 CTGGGCAGTTGATGCCATGGCTGACTGGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
          |||||||||||||||      |||||||||||||||||||   |||||||  |||||||||
Db    367 CTGGGCAGTTGATGCCGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGA 426

Qy    464 TATCATCTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGA 523
          ||||||||||||||||||||||||||||||||                  ||||||||||
Db    427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCCTACCTTACAGCAGTCAGTCTGGA 486

Qy    524 CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAAACTGCTGGCTGA 583
          |||||||||||||||                 ||||||||| ||||||| || |||||||
Db    487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy    584 AAAGGCAGTCAGTTCTATGTGGCCGTTCTGATCCCCTCCTCTGACTATACCTGACTTCAT 643
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    547 AAAGGTGGTCATGTTGGCGTTCTGGATCCCCTGGTCTGACTATTCCGACTTCAT 606

Qy    644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673
          |||||||||||||||||| ||| |||||
Db    607 CTTTGCCAACGTCAGTGAGGCGAGCAGATGACACAG 636
```

RESULT 12
ABZ68886
ID   ABZ68886 standard; cDNA; 1679 BP.
XX
AC   ABZ68886;
XX
DT   28-MAY-2003 (first entry)
XX
DE   Nucleotide sequence of human chemokine receptor CXCR4.
XX
KW   Human; chemokine receptor; CXCR4; viral infection; surface protein;
KW   respiratory virus infection; respiratory syncytial virus infection;
KW   RSV infection; bronchiolitis; bronchitis; pneumonia; asthma; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   89..1147
FT   /*tag= a

```
Qy    344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTT 403
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 GGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCATCACGTTCCTT 366

Qy    404 CTGGGCAGTTGATGCCATGGCTGACTGGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
          ||||||||||||||||      ||||||||||||||||||   ||||||| ||||||||||
Db    367 CTGGGCAGTTGATGCCGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGA 426

Qy    464 TATCATCTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGA 523
          |||||||||||||||||||||||||||||||||                 ||||||||||
Db    427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCCTACCTTACAGCAGTCAGTCTGGA 486

Qy    524 CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAAACTGCTGGCTGA 583
          |||||||||||||||                 ||||||||| ||||||| |||||||||
Db    487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy    584 AAAGGCAGTCAGTTCTATGTGGCCGTTCTGATCCCCTCCTCTGACTATACCTGACTTCAT 643
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    547 AAAGGTGGTCATGTTGGCGTTCTGGATCCCCTGGTCTGACTATTCCGACTTCAT 606

Qy    644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673
          ||||||||||||||||| ||| ||||||
Db    607 CTTTGCCAACGTCAGTGAGGCGAGCAGATGACACAG 636
```

RESULT 11
ABX74454
ID   ABX74454 standard; cDNA; 1679 BP.
XX
AC   ABX74454;
XX
DT   21-MAR-2003 (first entry)
XX
DE   Human cDNA sequence #26 up-regulated in CC-RCC patients.
XX
KW   Human; microarray; solid surface; immobilised probe; CC-RCC;
KW   differential expression profile; aggressive CC-RCC tumour type;
KW   non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
KW   gene expression profiling; tumour tissue; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200279411-A2.
XX
PD   10-OCT-2002.
XX
PF   29-MAR-2002; 2002WO-US009576.
XX
PR   29-MAR-2001; 2001US-0279411P.
XX
PA   (VAND-) VAN ANDEL INST.
XX
PI   Haab B, Rhodes D, Teh BT, Takashi M;
XX
DR   WPI; 2003-040679/03.
XX
PT   New microarray, comprising a matrix of cDNA probe from a set of probes
PT   immobilized to a solid surface in predetermined order, useful in the
PT   prognosis of patients with clear cell renal carcinoma.
XX
PS   Claim 35; Page 132-133; 179pp; English.
XX
CC   The present invention relates to a microarray comprising a matrix of at
CC   least one cDNA probe from a set of probes immobilised to a solid surface
CC   in a predetermined order, where a row of pixels corresponds to replicates
CC   of one distinct probe from the set. The probes are complementary to
CC   nucleic acid sequences that are expressed differentially in aggressive as
CC   compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
CC   and that hybridise to the probes under high stringency conditions. The
CC   microarray is useful for the prognosis of patients with CC-RCC, wherein
CC   aggressive and non-aggressive CC-RCC tumour types are characterised by
CC   differential expression profiles of genes that hybridise with one or more
CC   probes immobilised on the microarray. The arrays are useful for gene

FT   /product= "CXCR4"
XX
PN   WO200301414153-A2.
XX
PD   20-FEB-2003.
XX
PF   12-AUG-2002; 2002WO-CA001248.
XX
PR   10-AUG-2001; 2001US-0311088P.
XX
PA   (TOPI-) TOPIGEN PHARM INC.
XX
PI   Renzi P, Zemzoumi K;
XX
DR   WPI; 2003-256541/25.
DR   P-PSDB; ABP97733.
XX
PT   Modulating viral infection of a cell, for treating or preventing
PT   respiratory virus infections, bronchitis, pneumonia or asthma, by
PT   modulating a binding interaction between a cell chemokine-receptor and a
PT   surface protein of the virus.
XX
PS   Disclosure; Page 117-119; 120pp; English.
XX
CC   The present sequence encodes human chemokine receptor CXCR4. The
CC   specification describes a method for modulating viral infection of a
CC   cell. the method comprises modulating a binding interaction between a
CC   cell chemokine-receptor and a surface protein of the virus. The proviso
CC   is that the cell chemokine-receptor is not CX3CR1 and that the virus is
CC   not HIV. The method is useful for treating or preventing respiratory
CC   virus infection in vertebrates, more particularly respiratory syncytial
CC   virus (RSV) infections, and related diseases, e.g. bronchiolitis,
CC   bronchitis, pneumonia or asthma
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   65.2%;   Score 446.8; DB 7;   Length 1679;
Best Local Similarity   83.0%;   Pred. No. 7.7e-121;
Matches 523; Conservative   0; Mismatches 102; Indels   5;   Gaps   1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
           ||||||| |||||| ||||| |||| |||||||||||||| |||||| ||||||||||||
Db   12 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy   104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAAACTACTC 163
           || ||| ||| |||| |||| |||||| |||||| ||||||||||||||||| |
Db   72 GAGAACCACGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
           |||||||||| ||||||||| |||||||| |||||||||| |||| ||||||||||||||
Db   127 CGAGGAAATGGGCTCAGGGACCATGACTCCAATGACCGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy   224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTTACTTCATCATCTTCTTGACTGGCAT 283
           |||||| |||||| ||||||||||||||||||||| ||||||||||||||||||||||||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCATCTTCCATCATCTTCTTAACTGGCAT 246

Qy   284 AGTCGGGCAATGAGATTGGTGATCCTGGTCATGGTTACCAGAGAGAAGCTAAGGAGCATGAC 343
           ||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||
Db   247 TGTGGGGCAATGAGATTGGTCATCTCGGTCATGGTTACCAGAGAAGAAACTGAGAGAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCCTCTCTTTGTCATCACACTCCCCTT 403
           |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   307 GGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCCTCTCTTTGTCATCACGCTTCCCTT 366

Qy   404 CTCGGCAGTTGATGCCATGGCTGACTCGTGGTACTTTGGGAAATTTTGTCGTAAGGCTGTCCA 463
           ||||||||||||||||||||||||||| |||||||| |||||||| |||||||||||||||||
Db   367 CTCGGCAGTTGATGCCGTGGCAAACTGTAGTTTGGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
           |||||||| |||| |||||||||||||||||| |||| |||||||||||||||||||||
Db   427 TGTCATCTACACAGTCAACCTCAACTCTACACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
           |||||||||||| ||||| ||| |||||||||||||||| |||||||||| || |||||||
Db   487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGCCTCCCTGCTGATCAGCCCTGACTTCAT 546

Qy   584 AAAGGCAGTCTATGTGGGCGTCTGATCCGATCCCAGCCCTCCTGCTGACTATACCTGACTTCAT 643
           |||||| ||| |||||||| |||||||| |||| |||| ||||||||||||| ||||||||||
Db   547 AAAGGTGGTCTATGTTGGGCGTCTGGATCCTGATCCTGTGCCTCCTGTCGACTATTCCGGACTTCAT 606

Qy   644 CTTTGCCGACGTCAGCCAGCCAGGGGACATCAG 673
           ||||||||| ||||||||||||||||| |||||
Db   607 CTTTGCCAACGTCAGTCAGTGAGCGCAGCGAGATGACAG 636

RESULT 13
ABZ42642
ID   ABZ42642 standard; DNA; 1679 BP.
XX
AC   ABZ42642;
XX
DT   04-MAR-2003 (first entry)
XX
DE   Human CXC chemokine receptor 4 nucleotide SEQ ID NO:75.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
KW   Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW   osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW   graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW   psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW   mental retardation; epilepsy; asthma; tuberculosis; obesity; nusea;
KW   hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW   ulcer; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO20020261087-A2.
XX
PD   08-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US050107.
XX
PR   19-DEC-2000; 2000US-0257144P.
XX
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Burmer GC, Roush CL, Brown JP;
XX
DR   WPI; 2003-046718/04.
DR   P-PSDB; ABP81796.
XX
PT   New isolated antigenic peptides e.g., for G protein-coupled receptors
PT   (GPCR), useful for diagnosing and designing drugs for treating conditions
PT   in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT   autoimmune diseases.
XX
PS   Disclosure; Fig 1; 523pp; English.
XX
CC   The present invention describes antigenic peptides (I) comprising: (a)
CC   any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular G
CC   protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity or
CC   avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC   gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC   antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-related disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,

CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC   GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC   exemplification of the present invention
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

```
Query Match             65.2%;   Score 446.8;   DB 7;  Length 1679;
Best Local Similarity   83.0%;   Pred. No. 7.7e-121;
Matches  523;  Conservative   0;  Mismatches 102;  Indels   5;  Gaps   1;

Qy    44 GCAGGTGCAGGTAGCAGTGACCCTTCGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTG 103
         |||| ||| |||||| ||||| |||   ||||  |||||||||||| ||| |||||| |||
Db    12 GCGGCAGCAGGTAGCAAAGTGACCCGGAGGGCCTGAGTGCTCCAGTAGCCACCGCCATCTG 71

Qy   104 TAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATAACACTTCTGATAACTACTC 163
                |||     ||||| ||| || |||| || || |||||| ||| || || |||
Db    72 GAGAACCAGCCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAA 223
         |||| |||||||  ||||| ||| |||||||| || ||||||| ||| ||  ||||||||
Db   127 CGAGGAAATGGGCTTCAGGGGACTATGACTCCATGGAGGGAACCCTGTTTCCGTGAAGAAAA 186

Qy   224 CGTCCATTTCAATAGGATCTTCCTGCCACCACTACTTCATCATCTTCGACTGGCAT 283
         |  | |||||||||  ||||||||||||||| ||  || |||||||| ||||||||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCACCATCCATCACTCCATCATCTTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGAATGGATTGGTGATCCTCGGTCATGGGTTACCAGAAGAAGTAAGGACATGAC 343
         ||| |||||||  | ||| |||  ||||||||||||||||||||||||||  |||| |||||
Db   247 TGTGGGCAATGGAATTGGTCATCCTCGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGAC 306

Qy   344 GGACAAGTACCGGCGCTGACACCTCCCTTGTCATCACACTCCCCTT 403
         ||||||||| |||||||| ||||| |||||||||||| |  |||||
Db   307 GGACAAGTACAGGCGCTGCACCTGTCAGTGGCCGACCTCCTTGTCATCACGCGTTCCCTT 366

Qy   404 CTGGGCAGTTGATCCATGGCTGACTGGTACTTTGTGTAAGGCTGTCCA 463
         |||||||||||| |    ||||| ||||||||| ||| |||| |||||
Db   367 CTGGGCAGTTGATGCGCGTGCAAACTGGTACTTTGGGACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGCCCTTCATCAGCCTGGA 523
         | ||||||||| |||||||||||||||||||| |||||||| |||||||||||| |||||
Db   427 TGTCATCTACACAGTCAACCTCTACAGCAGCGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy   524 CCGGTACCTCGCCATTGTCCACGCCACCAAGGCCAAGGAAACTGCTGGCTGA 583
         |||||||||| ||||||||||||||||| |||||| ||||| |  ||||||
Db   487 CCGCTACCTGGCCATCGTCCACGCCACCAGTCAGAGGCGCAAGGAAGAGCTGTTGGCTGA 546

Qy   584 AAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGATCTACACCTGACTTCAT 643
         ||||| ||||||||| |||||||||||||| |||||||||| || |  |||||||||
Db   547 AAAGGTGGTCTATGTTGGGCGTCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCAT 606

Qy   644 CTTTGCCGACGTCAGCCAGCGGGGACATCAG 673
         ||||||| ||||||| |||| |  ||| |||
Db   607 CTTTGCCAACGTCAGCGAGCAGCAGATGACAG 636
```

RESULT 14
ADC98645
ID   ADC98645   standard; cDNA; 1679 BP.
XX
AC   ADC98645;
XX
DT   01-JAN-2004 (first entry)
XX
DE   Human CXC chemokine receptor 4 (CXCR4)-encoding cDNA.
XX
KW   Human; CXC chemokine receptor 4; CXCR4; CXCR4 inhibitor;

KW   small cell lung cancer; SCLC; cellular adhesion; cellular proliferation;
KW   metastasis; motility; morphological change; drug screening; monitoring;
KW   cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   89..1147
FT   /*tag= a
FT   /product= "Human CXCR4"
XX
PN   WO2003079020-A2.
XX
PD   25-SEP-2003.
XX
PF   20-MAR-2003; 2003WO-EP002916.
XX
PR   20-MAR-2002; 2002US-0366370P.
XX
PA   (DAND) DANA FARBER CANCER INST INC.
PA   (NOVS) NOVARTIS AG.
XX
PI   Salgia R;
XX
DR   WPI; 2003-757037/71.
DR   P-PSDB; ADC98646.
XX
PT   Use of a CXC chemokine receptor 4 (CXCR4) inhibitor, for modulating
PT   cellular adhesion or inhibiting proliferation, movement, or morphological
PT   change in a small cell lung cancer (SCLC) cell population, or for
PT   treating a subject having SCLC.
XX
PS   Disclosure; SEQ ID NO 1; 78pp; English.
XX
CC   The invention relates to methods and compositions for inhibiting small
CC   cell lung cancer (SCLC) proliferation and metastasis through modulation
CC   of the activity or expression of CXC chemokine receptor 4 (CXCR4), and
CC   optionally, that of the tyrosine kinase receptor c-Kit. Administration of
CC   a CXCR4 inhibitor modulates cellular adhesion and inhibits cellular
CC   proliferation, motility, and morphological changes in an SCLC cell
CC   population. The invention is based on the finding that CXCR4 is
CC   ubiquitously expressed, and c-Kit is variably expressed, in SCLC cells.
CC   The invention also discloses methods of screening for CXCR4 inhibitors;
CC   determining whether a sample of lung cancer cells expresses CXCR4 and is
CC   therefore amenable to CXCR4 inhibitor therapy; determining whether a
CC   sample of lung cancer cells coexpress CXCR4 and c-Kit is is therefore
CC   amenable to treatment with a combination of a CXCR4 inhibitor and a c-Kit
CC   inhibitor; and a method of assessing whether CXCR4 inhibitor therapy or
CC   CXCR4 inhibitor/c-Kit inhibitor combination therapy should be continued,
CC   by determining CXCR4 (and optionally c-Kit) activity in two or more
CC   tumour samples taken from an SCLC patient during the course of treatment.
CC   The methods of the invention are useful for treating SCLC, for screening
CC   for agents for the treatment of SCLC, for determining whether a SCLC from
CC   individual patients is suitable for treatment with agents of the
CC   invention, and for determining whether such treatment is effective and
CC   should be continued. The present sequence represents cDNA encoding human
CC   CXCR4, which can be used in methods of the invention.
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

```
Query Match             65.2%;   Score 446.8;   DB 9;  Length 1679;
Best Local Similarity   83.0%;   Pred. No. 7.7e-121;
Matches  523;  Conservative   0;  Mismatches 102;  Indels  102;  Gaps   1;

Qy    44 GCAGGTGCAGGTAGCAGTGACCCTTCGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTG 103
         |||| ||| |||||| ||||| |||   ||||  |||||||||||| ||| |||||| |||
Db    12 GCGGCAGCAGGTAGCAAAGTGACCCGGAGGGCCTGAGTGCTCCAGTAGCCACCGCCATCTG 71

Qy   104 TAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATAACACTTCTGATAACTACTC 163
                |||     ||||| ||| || |||| || || |||||| ||| || || |||
Db    72 GAGAACCAGCCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAA 223
```

CC The invention relates to a gene expression profile comprising one or more
CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
CC is a coronary artery endothelium, umbilical artery or vein endothelium,
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
CC endothelium, myometrium microvascular endothelium, keratinocyte
CC epithelium, bronchial epithelium, mammary epithelium, prostate
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
CC small airway epithelium, renal epithelium, umbilical artery smooth
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
CC osteoblasts or prostate stromal cell. The gene expression profile is used
CC for determining the level of RNA expression for a sample, determining the
CC phenotype of a cell and distinguishing cell types. The gene or a protein
CC expression profile is useful in identifying disease pathologies involving
CC alterations of gene expression. The assessment of expression profiles may
CC provide meaningful information with respect to tumour type and stage,
CC treatment methods, and prognosis. The gene or protein expression profile
CC may also be used for creating microarrays. The microarray is useful for
CC genetic and physical mapping of genomes, DNA sequencing, genetic or
CC medical diagnosis, genotyping of organisms, confirming cell or tissue
CC identifications and in identifying promising antibiotics, antiviral or
CC antifungal agents
XX
SQ Sequence 1711 BP; 423 A; 407 C; 373 G; 507 T; 0 U; 1 Other;

Query Match   65.2%;   Score 446.8;   DB 6;   Length 1711;
Best Local Similarity   83.0%;   Pred. No. 7.8e-121;
Matches 523;   Conservative   0;   Mismatches 102;   Indels   5;   Gaps   1;

Qy   44 GCAGGTGCAGTGAGCAGTGACCCTGAGCGTTGGTGCTCCGGTAACCACACGGCTG 103
Db   47 GCGGCAGCAGTGAGCAAAGTGACGGCCGAGGCCTCCAGTGCTCCAGTAGCCACCGCCATCTG 106

Qy   104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
Db   107 GAGAACCAGCCGGTTACCATGGA----GGGGATCAGTATATACACTTCAGATAACTACACAC 161

Qy   164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGAACCCTGCTTCGGGATGAAAA 223
Db   162 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 221

Qy   224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
Db   222 TGCTAATTCAATAAAATCTTCCTGCCCACCATCTCATCCATCTTCTTAACTGGCAT 281

Qy   284 AGTCGGCAATGGTGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC 343
Db   282 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGAC 341

Qy   344 GGACAAGTACCGCTGCACCTGTCAGTGGCTGACCCTCCTTGTCATCACACTCCCTT 403
Db   342 GGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCCTCCTTGTCATCACGCTTCCCTT 401

Qy   404 CTGGGCAGTTGATGCCATGGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
Db   402 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 461

Qy   464 TATCATCTACACTGTCAACCTCTACAGCAGCGGTTCATCCTGGCCTTCATCAGCCTGGA 523
Db   462 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTCGA 521

Qy   524 CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGA 583
Db   522 CCGCTACCTGGCCATCGTCCACGCGTCCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 581

Qy   584 AAAGGCAGTCTATGTGGGGCGTCTGGATCCCAGGCCCTCCTCCTGACTATACCTGACTTCAT 643
Db   582 AAAGGTGGTCTATGTTGGCGGTCTGGATCCCCCTCCCTGACTATTCCCGACTTCAT 641

Qy   644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673

Db   127 CGAGGAAATGGGCTCAGGGGACTATGACTCCACCATCTACTTCATCATCTTCTTGACTGGCAT 186
Db   127 CGAGGAAATGGGCTCAGGGGACTATGACTCCACCATCTACTTCATCATCTTCTTGACTGGCAT 186
Qy   224 CGTCCATTCAATAGGATCTTCCTGCCCACCATCTCATCCATCTTCTTAACTGGCAT 246
Db   187 TGCTAATTCAATAAAATCTTCCTGCCCACCATCTCATCCATCTTCTTAACTGGCAT 246
Qy   284 AGTCGGCAATGGTGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC 343
Db   247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGAC 306
Qy   344 GGACAAGTACCGCTGCACCTGTCAGTGGCTGACCCTCCTTGTCATCACACTCCCTT 403
Db   307 GGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCCTCCTTGTCATCACGCTTCCCTT 366
Qy   404 CTGGGCAGTTGATGCCATGGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
Db   367 CTGGGCAGTTGATGCCGTGGCAAACTGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 426
Qy   464 TATCATCTACACTGTCAACCTCTACAGCAGCGGTTCATCCTGGCCTTCATCAGCCTGGA 523
Db   427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTCGGA 486
Qy   524 CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGA 583
Db   487 CCGCTACCTGGCCATCGTCCACGCGTCCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546
Qy   584 AAAGGCAGTCTATGTGGGGCGTCTGGATCCCAGGCCCTCCTCCTGACTATACCTGACTTCAT 643
Db   547 AAAGGTGGTCTATGTTGGCGGTCTGGATCCCCCTCCCTGACTATTCCCGACTTCAT 606
Qy   644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673
Db   607 CTTTGCCAACGTCAGTGAGGCAGATGACAG 636

RESULT 15
ABZ35348
ID   ABZ35348 standard; cDNA; 1711 BP.
XX
AC   ABZ35348;
XX
DT   05-FEB-2003 (first entry)
XX
DE   Human gene expression profile polynucleotide SEQ ID NO 459.
XX
KW   Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
KW   bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
KW   tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
KW   gene expression; gene; 88.
XX
OS   Homo sapiens.
XX
PN   WO200274979-A2.
XX
PD   26-SEP-2002.
XX
PF   20-MAR-2002; 2002WO-US008455.
XX
PR   20-MAR-2001; 2001US-0276947P.
XX
PA   (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI   Wan J; Wang Y;
XX
DR   WPI; 2002-740862/80.
XX
PT   New gene expression profile generated from primary, endothelial,
PT   epithelial, and muscle cell types, useful for identifying disease
PT   pathologies involving alterations of gene expression, e.g. cancer.
XX
PS   Disclosure; Page 609; 850pp; English.
XX

Db   642 CTTTGCCAACGTCAGTGAGGCAGATGACAG 671

Search completed: May 17, 2004, 06:37:07
Job time : 295.749 secs

```
                  GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   May 17, 2004, 05:52:09 ; Search time 54.8443 Seconds
                                (without alignments)
                           6931.281 Million cell updates/sec

Title:          US-09-367-052-5
Perfect score:  685
Sequence:       1 ccatcctaatacgactcact.........gacatcagtcaggggatga 685

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:      1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued_Patents_NA:*
        1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
        2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
        3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:**
        4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
        5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
        6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	446.8	65.2	1679	4	US-09-517-605-14	Sequence 14, Appl
2	445.8	65.1	1664	4	US-09-582-224A-5	Sequence 5, Appli
3	445.8	65.1	1664	4	US-09-023-655-1213	Sequence 1213, Ap
4	439.2	64.1	1737	1	US-08-202-056-4	Sequence 4, Appli
5	439.2	64.1	1737	1	US-09-076-093A-3	Sequence 3, Appli
6	439.2	64.1	1737	1	US-08-701-265-3	Sequence 3, Appli
7	439.2	64.1	1737	2	US-08-284-586-3	Sequence 3, Appli
8	439.2	64.1	1737	2	US-08-805-478-3	Sequence 3, Appli
9	439.2	64.1	1737	2	US-08-802-627A-3	Sequence 3, Appli
10	439.2	64.1	1737	2	US-08-801-238-3	Sequence 3, Appli
11	439.2	64.1	1737	2	US-08-801-228-3	Sequence 3, Appli
12	439.2	64.1	1737	3	US-09-104-296-3	Sequence 3, Appli
13	439.2	64.1	1737	5	PCT-US94-06380-2	Sequence 2, Appli
14	430.8	62.9	1225	4	US-09-016-434-1235	Sequence 1235, Ap
15	428	62.5	1317	1	US-08-153-848-45	Sequence 45, Appl
16	428	62.5	1317	4	US-09-299-843A-45	Sequence 45, Appl
17	428	62.5	1317	4	US-09-088-337B-45	Sequence 45, Appl
18	428	62.5	1317	5	PCT-US93-11153-45	Sequence 45, Appl
19	152.4	22.2	1119	4	US-09-170-496D-65	Sequence 65, Appl
20	152.4	22.2	1119	4	US-09-170-496D-199	Sequence 199, App
21	152.4	22.2	1679	1	US-08-202-056-6	Sequence 6, Appli
22	152.4	22.2	1679	1	US-09-076-093A-5	Sequence 5, Appli
23	152.4	22.2	1679	1	US-08-701-265-5	Sequence 5, Appli
24	152.4	22.2	1679	2	US-08-284-586-5	Sequence 5, Appli
25	152.4	22.2	1679	2	US-08-805-478-5	Sequence 5, Appli
26	152.4	22.2	1679	2	US-08-802-627A-5	Sequence 5, Appli
27	152.4	22.2	1679	2	US-08-801-238-5	Sequence 5, Appli
28	152.4	22.2	1679	2	US-08-801-228-5	Sequence 5, Appli
29	152.4	22.2	1679	3	US-09-104-296-5	Sequence 5, Appli
30	152.4	22.2	1679	5	PCT-US94-06380-3	Sequence 3, Appli
31	152.4	22.2	2818	3	US-08-982-493-7	Sequence 7, Appli
32	152.4	22.2	2818	3	US-08-628-655-1	Sequence 1, Appli
33	141.2	20.6	2085	4	US-09-299-843A-65	Sequence 65, Appl
34	141.2	20.6	2085	4	US-09-088-337B-65	Sequence 65, Appl
35	139.4	20.4	1029	3	US-09-116-498-5	Sequence 5, Appli
36	138.8	20.3	2751	1	US-08-153-848-23	Sequence 23, Appl
37	138.8	20.3	2751	4	US-09-299-843A-23	Sequence 23, Appl
38	138.8	20.3	2751	4	US-09-088-337B-23	Sequence 23, Appl
39	138.8	20.3	2751	5	PCT-US93-11153-23	Sequence 23, Appl
40	137.8	20.1	1037	3	US-09-116-498-3	Sequence 3, Appli
41	131.8	19.2	1107	4	US-09-170-496D-19	Sequence 19, Appl
42	131.8	19.2	1107	4	US-09-170-496D-173	Sequence 173, App
43	131.8	19.2	1293	4	US-09-016-434-1052	Sequence 1052, Ap
44	131.8	19.2	1670	3	US-08-709-838-1	Sequence 1, Appli
45	131.8	19.2	1670	3	US-08-829-839-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-517-605-14
; Sequence 14, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
;  APPLICANT: Littman, Dan R.
;  APPLICANT: Kwon, Douglas S.
;  APPLICANT: van Kooyk, Yvette
;  APPLICANT: Geijtenbeck, Tneo
;  TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
;  TITLE OF INVENTION: CELLS
;  FILE REFERENCE: 1049-1-017
;  CURRENT APPLICATION NUMBER: US/09/517,605
;  CURRENT FILING DATE: 2000-03-02
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;  LENGTH: 1679
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-517-605-14

Query Match             65.2%;  Score 446.8;  DB 4;  Length 1679;
Best Local Similarity   83.0%;  Pred. No. 1e-116;
Matches   523;  Conservative   0;  Mismatches  102;  Indels   5;  Gaps   1;

Qy     44  GCAGGTGCAGTAGCAGTGACCCTCTGGAGGCGTTGGTGCTCCGGTAACCACCACCGGCTG 103
           ||||||||||||||  |||||||||||||||||||||||||| ||  ||||||| ||||||
Db     12  GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy    104  TAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATCACTTCTGATAACTACTC 163
           ||||  |||| |||||||||||||||         ||||||||||||||||||||||||
Db     72  GAGAACCAGCGGGTTACCATGGA-----GGGGATCAGTAGTATACACTTCAGATAACTACAC 126

Qy    164  TGAAGAGTGGGGTCTGGAGACTATGACTCAACAACAGGAACCCGTTCCGGGATGAAAA 223
           ||||||   |||||||||||||||||||       |||||||| ||| ||||||||||
Db    127  CGAGGAAATGGGCTCAGGGGACTATGACTATGACTCCATGAAGAACCCTGTTTCCGTGAAGAAAA 186

Qy    224  CGTCCATTTCAATAGGATCTTCCTGCCACCATCTACTTCTTGACTGGCAT 283
           |||||||||||||   |||||||||||||||||   ||| || |||||||
Db    187  TGCTAATTTCAATAAAATCTTCCTGCCACCATCTCCATCATCTTCTAACTGGCAT 246

Qy    284  AGTGCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC 343
           |||||||||||||| |||| ||||||||||||||||||||||||||  |||  ||||||||
Db    247  TGTGGGCCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAGAGCATGAC 306

Qy    344  GGACAAGTACCGGCTGCACCTGGTCAGTGGCTGACCTCCTTTGTCACTCACTCCCCTT 403
           |||||||||   ||||||||||  ||||||||| |||| || ||||| ||| || |||||
Db    307  GGACAAGTACAGCCTGCACCTGTCAGTGGCCGACACCTCTCCTTGTCATCACGCTTCCCTT 366
```

Qy 405 TGGGCAGTTGATGCCATGCTGCTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCAT 464
              |||||||||||||| |||||||||||||||||| ||||| | |||||||||||
Db 356 TGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCAT 415

Qy 465 ATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGAC 524
              ||||||||| |||||||||||||||||| || ||||||||||||||||||| |||||
Db 416 GTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGAC 475

Qy 525 CGGTACCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAA 584
              || ||||||||| ||||||||||||||||||| ||||||||||| ||| ||||
Db 476 CGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGGAGCTGTTGGCTGAA 535

Qy 585 AAGGCAGTCTATGTGGGCGTCTTGGATCCCAGCCCTCCTCCGACTATACCTGACTTCATC 644
              |||| |||||||| ||||||||||||||| ||||| |||||||| |||||||||
Db 536 AAGGTGGTCTATGTTGGCGTCTTGGATCCCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCATC 595

Qy 645 TTTGCCGACGTCAGCCAGGGGGACATCAG 673
              |||||| |||||| ||| |||| ||
Db 596 TTTGCCAACGTCAGTGAGGCAGATGACAG 624

RESULT 3
US-09-023-655-1213
; Sequence 1213, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g219868
US-09-023-655-1213

Query Match   65.1%;   Score 445.8; DB 4; Length 1664;
Best Local Similarity   83.0%;   Pred. No. 1.9e-116;
Matches 522;   Conservative 0; Mismatches 102; Indels 5; Gaps 1;

Qy 404 CTGGGCAGTTGATGCCATGCTGCTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
              |||||||||||||||| ||||||||||||||||| ||||| | ||||||||
Db 367 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy 464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
              | ||||||||| |||||||||||||||||| || ||||||||||||||||||| |||||
Db 427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy 524 CCGGTACCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
              ||| ||||||||| ||||||||||||||||||| ||||||||||| ||| ||||
Db 487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGGAGCTGTTGGCTGA 546

Qy 584 AAAGGCAGTCTATGTGGGCGTCTTGGATCCCAGCCCTCCTCCGACTATACCTGACTTCAT 643
              ||||| |||||||| |||||||||||||||| ||||| |||||||| |||||||||
Db 547 AAAGGTGGTCTATGTTGGCGTCTTGGATCCCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCAT 606

Qy 644 CTTTGCCGACGTCAGCCAGGGGGACATCAG 673
              |||||||| |||||| ||| |||| |
Db 607 CTTTGCCAACGTCAGTGAGGCAGATGACAG 636

RESULT 2
US-09-582-224A-5
; Sequence 5, Application US/09582224A
; Patent No. 6429308
; GENERAL INFORMATION:
; APPLICANT: IIJIMA, Osamu
; APPLICANT: GOTO, Takeshi
; APPLICANT: SHIMADA, Takashi
; TITLE OF INVENTION: HIV Infection Inhibitors
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/582,224A
; CURRENT FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: PCT/JP99/06534
; PRIOR FILING DATE: 1999/11/24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cDNA of CXCR4
US-09-582-224A-5

Query Match   65.1%;   Score 445.8; DB 4; Length 1664;
Best Local Similarity   83.0%;   Pred. No. 1.9e-116;
Matches 522;   Conservative 0; Mismatches 102; Indels 5; Gaps 1;

Qy   45 CAGGTGCAGGTAGCAGTGACCCTCTGAGGCCGTTGGTGCTCCGGTAACCACCACGGCTGT 104
              || ||||||||||| ||| ||| ||||||||| |||| ||| ||||||||||||
Db   1 CGGCAGCAGGTAGCAAAGTGACGCGCCGAGGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGG 60

Qy 105 AGAGCCAGGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCT 164
              |||| |||| | ||| || ||||| |||| ||||||||||||| ||||| |||||
Db   61 AGAACCAGCCGGTTACCATGGA-----GGGGATCAGTCAGTATATACACTTCAGATAACTACACC 115

Qy 165 GAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAC 224
              || ||| | |||||| ||| |||||| ||||||||||| ||||| |||||||||
Db 116 GAGGAAAATGGGCTCAGGGGACTCATGAACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAAT 175

Qy 225 GTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTCGGCATA 284
              || ||| |||||| | ||| |||||| ||||||| |||||||||||| ||| ||||||
Db 176 GCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTTCTTAACTGGCATT 235

Qy 285 GTCGGGCAATGGATTGGTGATCCTGGTCATGGTTACCAGAAGAAGCTAAGGAGCATGACG 344
              |||| |||||| |||||| |||||| || ||||||||||||| ||| | ||||||||
Db 236 GTGGGCAATGGATTGGTCATCCTGGTCATGGTTACCAGAAGAAACTGAGAAGCATGACG 295

Qy 345 GACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTC 404
              |||||||||| |||||||||||||||||| ||||||||||| | |||||| |||||
Db 296 GACAAGTACAGCGGCTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCACGCTTCCCTTC 355

Query Match   64.1%;   Score 439.2;   DB 1;   Length 1737;
Best Local Similarity 82.7%;   Pred. No. 1.4e-114;
Matches 516;   Conservative   0;   Mismatches 103;   Indels   5;   Gaps   1;

```
Qy  50 GCAGGTAGCAGTGACCCTGAGGCGTTGGTGTCCGGTAACCACCACGGCGTGTAGAGC 109
          ||| ||||||||| ||||  |  ||| || ||||  |||||||| ||||| |||||
Db  20 GCGCGGCGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy 110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
             ||||| ||||| |     ||||| ||||||||||||| ||||||||| || ||||
Db  80 CAGCGGGTTACCATGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGA 134

Qy 170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
          ||||| ||||| |||||||||||||| ||||||||||| |||||| |||||| | ||
Db 135 AATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy 230 TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGG 289
          ||||||| |||| ||||||||||||||||||||||||||||||| |||||||| || |
Db 195 TTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTTCTTAACTGGCATTGTGGG 254

Qy 290 CAATGGATTGGTGATCCTGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAA 349
          ||||||||||||| ||||| |||||||||||||||||| |||| |||||||||||||||
Db 255 CAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGACGGACAA 314

Qy 350 GTACCGGCTGCACCTGTCAGTGCTGCTGACCCTCCTCTTGTCATCACACTCCCCTTCTCGGGC 409
          ||| ||||| ||||||||||||||| ||||||||| | |||||||| ||| ||||||||||
Db 315 GTACAGGCTGCACCTGTCAGTGCTGACCCTCCTTTGTCATCACGCTTCCCTTCTCTCTGGC 374

Qy 410 AGTTGATGCCATGGCCATGGTACTTTGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
          |||| |||| |||| ||||||| ||||||| | |||  || |||||| |||| ||||
Db 375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy 470 CTACACTGTCAACCTCTACAGCACCTTCATCCTGGCCTTCATCAGCGCCTGGACCGGTA 529
          |||||||||||||||||||| || |||||||| | ||||||||| |||||| ||| |||
Db 435 CTACACAGTCAACCTCTACAGCAGTGTCCTCATCTGGCCTTCATCTCGGCAACCCCTA 494

Qy 530 CCTCGCCATTGTCCACGCCACCAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGC 589
          || ||||||||||| ||||||||||||||| ||||||||||| |||||||||  |||||
Db 495 CCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTGTGGCTGAAAAGGT 554

Qy 590 AGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGGACTATACCTGACTTCATCTTTGC 649
          |||| ||||||| |||||||||||||||| | ||| |||| |||| |||||||||||||
Db 555 GGTCTATGTGGGCGTCTGGATCCCCTGCCCTCCTGCTGACTATTCCCGACTTCATCTTTGC 614

Qy 650 CGACGTCAGCCAGGGGGACATCAG 673
          ||| ||||||||| |||||| |||
Db 615 CAGCGTCAGTGAGGCAGATGACAG 638
```

RESULT 5
US-08-076-093A-3
; Sequence 3, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:

```
Qy  45 CAGGTGCAGGTAGCAGTGACCCTGAGGCGGTTGGTGTCCGGTAACCACCACGGCGTGT 104
          | ||||||||| ||||  |  ||| || ||||  |||||||| ||||| ||||
Db   1 CGGCGGCAGGTAGCAGAAGTGACGGCCGAGGGCCTGAGTGCTGCCAGTAGCCACCGCATCTGG 60

Qy 105 AGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCT 164
          |     ||||| ||||| |     ||||| ||||||||||||| ||||||||| ||
Db  61 AGAACCAGCGGGTTACCATGA-----GGGGATCAGTATATACACTTCAGATAACTACACACC 115

Qy 165 GAAGAAGTCGGGGTTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAC 224
          |||| ||||| ||||| |||||||||||||| |||||||||| |||||| |||||| |
Db 116 GAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAAT 175

Qy 225 GTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATA 284
          | |  ||||||| |||| ||||||||||||||||||||||||||||||| ||||||||
Db 176 GCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTTCTTAACTGGCATT 235

Qy 285 GTCGGCAATGGATTGGTGATCCTGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACG 344
          || |||||||||||||| ||||| |||||||||||||||||| |||| |||||||||||
Db 236 GTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGACG 295

Qy 345 GACAAGTACCGGCTGCACCTGTCAGTGCTGCTGACCCTCCTCTTGTCATCACACTCCCCTTC 404
          ||||| ||||| ||||||||||||||| ||||||||| | |||||||| ||| ||||||
Db 296 GACAAGTACAGGCTGCACCTGTCAGTGCTGACCCTCCTTTGTCATCACGCTTCCCTTC 355

Qy 405 TGGGCAGTTGATGCCATGGCCATGGTACTTTGGAAATTTTGTGTAAGGCTGTCAT 464
          ||| ||||| |||| |||| ||||||| |||||| | |||  || ||||| ||
Db 356 TCTGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCAT 415

Qy 465 ATCATCTACACTGTCAACCTCTACAGCACCTTCATCCTGGCCTTCATCAGCGCCTGGAC 524
          ||||||||||||||||||||||||| || |||||||| | ||||||||| ||||||
Db 416 GTCATCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCTGGCCTTCATCTCGGCTCTGAC 475

Qy 525 CGGTACCTCGCCATTGTCCACGCCACCAGTCAAAGGCCAAGGAAACTGCTGGCTGAA 584
          | ||||| ||||||||||| ||||||||||| ||||||||||||| ||| |||||||
Db 476 CCCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTGTGGCTGAA 535

Qy 585 AAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGGACTATACCTGACTTCATC 644
          ||| |||| ||||||| |||||||||||| | |||| ||| |||| |||| |||||||
Db 536 AAGGTGGTCTATGTGGGCGTCTGGATCCCCTGCCCTCCTGCTGACTATTCCCGACTTCATC 595

Qy 645 TTTGCCGACGTCAGCCAGGGGGACATCAG 673
          |||||| || ||||||||| |||||
Db 596 TTTGCCAGCGTCAGTGAGGCAGATGACAG 624
```

Query Match   64.1%;   Score 439.2; DB 1; Length 1737;
Best Local Similarity 82.7%;   Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

```
Qy   50 GCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC 109
          ||||||  || || ||||||||||| |||||||| |||||||||||||||||||||||
Db   20 GCGCGGCGCAAAGTGACGCGCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy  110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
          ||    ||   ||  ||||  |||||||||||||||||||| |||||||| |||  |||
Db   80 CAGCGGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTAACTACACCGAGGA 134

Qy  170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
          || ||| |||||||| |||||||||||||||||||||||||||||||||||||| ||||
Db  135 AATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAAATGCTAA 194

Qy  230 TTTCAATAGGATCTCCTGCCCACCATCCACTCACTCCATCATCTCTTAACTGGCATTGTCGG 289
          ||||||||| ||||||||||||||||||||| || ||||||||||||||||||||||||
Db  195 TTTCAATAAAATCTTCCTGCCCACCACCACTCCATCATCATCTCTTAACTGGCATTGTGGG 254

Qy  290 CAATGGATTGGTGATCTGGTCATGGCTTACCAGAGAAGCTAAGGAGCATGACGGACAA 349
          ||||||||| |||||||||||||||||||||||||||||||| ||||||||||||||
Db  255 CAATGGATTGGTCATCCTGGTCATGGGTTACCAGGAGAAGCATGAGAAGCATGACGGACAA 314

Qy  350 GTACCGGCTGCACCTGTGACCTGTGGCTGACCTCCCTTTGTCATCACACTCCCCTTCTGGGC 409
          ||||||||| |||||||||||||| ||||||||||||||||||||||||||||| ||||||
Db  315 GTACAGGCTGCACCTGTCAGTGAGTGGCCGACCTCCCTTTGTCATCACGCTTCCCTTCTGGGC 374

Qy  410 AGTTGATGCCATGGCTGACTGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCATATCAT 469
```

```
Db  375 AGTTGATGCCGTGCCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCATGTCAT 434
          ||||||||| || |  ||  |||||  ||| |||||| |||| ||||| |||| |||||

Qy  470 CTACACTGTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCTCGGCCTTCATCAGCCTGGACCGGTA 529
          |||||| |||||||| ||||||||||| |||||||||||||||||| || |||||||||||| |

Db  435 CTACACAGTCAACCTCTACAGCAGTGTCCTCAGCCTTCATCCTCGGCCTTCATCAGTCTCGGACCGCTA 494

Qy  530 CCTGCCATTGTCACGCCACCACCACAGTCAAAGGCAAGGAAACTGCTGGCTGAAAAAGGC 589
          ||||| ||||||||||||||||||||| |||  |||||| |||||  ||||||||| |||

Db  495 CCTGGCCATCGTCACGCCACCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAAGGT 554

Qy  590 AGTCTATGTGGCGGCTGCGGTCGTGGGTGTCCCAGCCCTCCCTGACTCCTGACTTCATCTTTGC 649
          |||||||| ||||| ||| ||| || ||||||||||||||| ||||| ||||| |||||||||||

Db  555 GGTCTATGTTGGCGGTCTGGAGATCCCTGGATCCCCTGCCTTCCCTGCTGACTATTCCCGACTTCATCTTTGC 614

Qy  650 CGAGTCAGCCAGGGGGACATCAG 673
          || |||||||  ||| || |||

Db  615 CAACGTCAGTGAGGCAGATGACAG 638
```

Query Match   64.1%;   Score 439.2; DB 1; Length 1737;
Best Local Similarity 82.7%;   Pred. No. 1.4e-114;

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 nucleotides
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-284-586-3

Query Match 64.1%; Score 439.2; DB 2; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

```
Qy   50 GCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCGTGTAGAGC 109
           ||| ||         ||||  |||||  ||||| ||  |||||| || |   ||||||
Db   20 GCGCGGGCGCAAAGTGACGCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCGGAGAAC  79

Qy  110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCGATAACTACTCTGAAGA 169
           ||| |  |||||||      ||||| |||||||||||||||  ||||||||| ||||
Db   80 CAGCGGGTTACCATGGA----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134

Qy  170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
           | ||||| | ||| |||||||||||  ||| |||||||| ||||| ||||| | | |
Db  135 AATGGGCTCAGGGGACTATGACTCCATGAAGAGGAACCCTGTTTCCGTGAAGAAATGCTAA 194

Qy  230 TTTCAATAGGATCTTCCTGCCCACCATTCATCTTCTTGACTGGCATAGTCGG 289
          |||||||  | |||||||||||||||| | ||||||| ||||||| || |||
Db  195 TTTCAATAAAATCTTCCTGCCCACCATCTTCTTAACTGGCATTGTGGG 254

Qy  290 CAATGGATTGGTGATCCTGTCATGGGTTACCAGAAGAGCTAAGGAGCATGACGGACAA 349
          |||||||||||| ||||||||||||||||||||||||   ||| ||||||||||| ||
Db  255 CAATGGATTGGTCATCCTGTCATGGGTTACCAGAAGAAACTGAGAAGCATGACGGACA 314

Qy  350 GTACCGGCTGCACCTGTCAGTGGCTGACCTGCTGACCTCCTTTGTCATCACCACTCCCCTTCTCGGGC 409
          |||| ||||||||||||||||||||||||||| |||||||||||||||||| ||| ||||||||||||
Db  315 GTACAGGCTGCACCTGTCAGTGGCTGACCTGCCGACCTCCTTTGTCATCACGCTTCCCTTCTCGGGC 374

Qy  410 AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCAT 469
          |||||||||| |||||||| ||||| |||||||| |||||| |||||| | ||||| |||
Db  375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTTCCATGTCAT 434

Qy  470 CTACACTGTCAACCTTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTA 529
          || ||| |||||||||||| ||||||| |   |||||| |||||||| ||||||||||| |
Db  435 CTACACAGTCAACCTTCTACAGCAGTGTCTTCCTCATCAGTCAGTCTGGACCGCTA 494

Qy  530 CCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGC 589
          ||||||||| |||||||| |||| ||||||||||||||||||||||  || ||||||||
Db  495 CCTGGCCATGTCCACGCCACCAACAGTCAGGGCCAAGGAAGCTGTTGGCTGAAAAGGT 554

Qy  590 AGTCTATGTGGGCGGTCTGGATCCCAGCCCCTCCTGACTATACCTGACTTCATCTTTGC 649
          |||| |||| |||| | |||||| ||||  |||||| | ||||| |||||||||| |||
Db  555 GGTCTATGTTGGCGGTCTGGATCCTGATCCCCTGCTGCTATTCCCGACTTCATCTTTGC 614

Qy  650 CGACGTCAGCCAGGGGGACATCAG 673
          | | |||||| ||| || ||| ||
Db  615 CAACCTCAGTGAGGCAGATGACAG 638
```

Qy 290 CAATGGATTGGTGATCCTCGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGACGGACAA 349
         ||||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||
Db 255 CAATGGATTGGTCATCCTCGTCATGGGTTACCAGAAGAAGCTGAGGAGCATGACGGACA 314

Qy 350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGC 409
         |||| ||||||||||||||||||||| | |||||||||||||||| || ||| |||||
Db 315 GTACAGGCTGCACCTGTCAGTGGCTGGCCGACCTCCTCTTTGTCATCACGCTTCCCTTCTGGGC 374

Qy 410 AGTTGATGCCATGGCCATGGTACTTGGGAAATTTTGTGTAGGCTGTCCATATCAT 469
         |||||||||| |||| | |||||| |||| | | | ||||||| |||||
Db 375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy 470 CTACACTGTCAACCTCTACAGCACAGCGTTCATCCTGGCCTTCATCAGCCTGACCGGGTA 529
         |||| |||||||||||||| ||||||| ||| ||| ||||||||| |||| ||||||
Db 435 CTACACAGTCAACCTTCTACAGCAGTGTCCTTCCTCTGGCCTTCATCCATCTGGACCGCTA 494

Qy 530 CCTCGCCATTGTCCACGCCACCAGTCAAAGGCCAAGAAACTGCTGCTGAAAAGGC 589
         ||| |||||| |||| |||| |||| || |||||||| ||||| |||||||||||
Db 495 CCTGCCATCGTCCACGCCACCAGTCAGAGGCCAAGAAGCTGTTGCTGCTGAAAAGT 554

Qy 590 AGTTCATGTGGGCGTCTGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGC 649
         |||||||| ||| ||| || |||| |||| ||| ||||||| |||||| |||||||||
Db 555 GGTCTATGTTGGGCGTCTGGGCTCCCCGCCCCTCCTGCTGACTATTCCCGACTTCATCTTTGC 614

Qy 650 CGACGTCAGCCAGGGGGGACATCAG 673
         | |||||||| || ||||| ||||
Db 615 CAACGTCAGTGAGGCAGATGACAG 638

RESULT 8
US-08-805-478-3
; Sequence 3, Application US/08805478
; Patent No. 5874543
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/805,478
;   FILING DATE: 25-Feb-1997
;   CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/284586
;   FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/076093
;   FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/810782
;   FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B.
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: P0706P2P1C1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1737 base pairs
;   TYPE: Nucleic Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Linear
US-08-805-478-3

Query Match   64.1%;   Score 439.2; DB 2; Length 1737;
Best Local Similarity   82.7%;   Pred. No. 1.4e-114;
Matches   516; Conservative   0; Mismatches 103; Indels   5; Gaps   1;

Qy 50 GCAGGTAGCAGTGACCCCTCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTGTGAGC 109
          || || ||| ||||||||||||| | | | | |||| ||| ||||| || |||
Db 20 GCGGCGGCCGCAAAGTGACGCCGAGGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAAC 79

Qy 110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
          ||||| | || || | |||||| || |||||||||||| ||||| |||||||||| |
Db 80 CAGCGGTTACCATGA-----GGGGATCAGTATATACACTTCAGATAACACTACACCGAGGA 134

Qy 170 AGTTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCCGGGATGAAAACGTCCA 229
          | ||||||| | | || |||||||||||| |||||||| |||||| | | |||||
Db 135 AATGGGCTTCAGGGGACTATGACTTCCATGAAGGAACCCTGTTTCCGGTGAAGAAAATGCTAA 194

Qy 230 TTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGCATAGTCGG 289
          |||||| | ||||||| |||||| ||||| ||| |||| |||| |||| |||| |||||
Db 195 TTTCAATAAAATCTTCCTGCCCCACCATCTCCATCATCTTCTTAACTGGCATTGTGGGG 254

RESULT 9
US-08-802-627A-3
; Sequence 3, Application US/08802627A
; Patent No. 5892017
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/802,627A
;   FILING DATE: 19-Feb-1997
;   CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/284586
;   FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/076093
;   FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/810782
;   FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B.
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: P0706P2P1D2
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:

LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-802-627A-3

Query Match 64.1%; Score 439.2; DB 2; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

```
Qy   50 GCAGGTAGCAGTGACCCCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTGTAGAGC 109
          |||| |||||||||| |||   ||| |||||||||||| ||||||  |||||||  ||  
Db   20 GCGGCGGCGCAAAGTGACGGCCGAGGGCCTGAGTGCTCCAGTAGCGTCCCAGTAGGAAC   79

Qy  110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
          |||    ||||||||||     ||||  |||||||||||||| |||||||||  ||| ||
Db   80 CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134

Qy  170 AGTCGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
          | ||||||| ||||||||||||||||  |||||||||||| ||||| ||||||| ||  
Db  135 AATGGGCTCAGGGGACTATGACTCCATGAAGAAGAACCCTGTTTCCGTGAAGAAATGCTAA 194

Qy  230 TTTCAATAGGATCTTCCTGCCCACCATCACTTCTTTCATCATCATCTTCTTAACTGGCATAGTCGG 289
          ||||||| |||| ||||||||||||||||| |||||||||||| |||||| ||| |||||||||||
Db  195 TTTCAATAAAATCTTCCTGCCCACCATCACTTCTTTCATCATCATCTTCTTAACTGGCATTGTGGG 254

Qy  290 CAATGGATTGGTGATCCTGGTTACCAGAAGAAGCTAAGGAGCATAGGACGGACAA 349
          ||||| |||||| |||||| ||||||| |||||| ||||| |||| |||| |||
Db  255 CAATGGATTGGTCATCCTGGTTACCAGAAGAAACTGAGAAGCATGAGCATGACGACAA 314

Qy  350 GTACCGGCTGCCACCTGTCAGTGCTGGCTGACCCTCCTTGTCATCATCACACTCCCCTTCTGGGC 409
          ||||||| |||||| |||||| |||||  ||||| ||| |||||||||||  |||||| ||||||
Db  315 GTACAGGCTGCCACCTGTCAGTGGGCCGAGACTTGGTAAGGCTGTCCATATCAT 374

Qy  410 AGTTGATGCCATGGCCTGACTGGTACTTTGGGAAATTTTCTGTGTAAGGCTGTCCATATCAT 469
          |||||||||||||||| |||||||||||||||||||||||||| ||||| ||||||||||||
Db  375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTTCCATGTCAT 434

Qy  470 CTACACTGTCAACCTTCTACAGCAGCTTCATCCTGGCCTTCATCAGCCTGGACCCGGTA 529
          ||||||||||||||||| |||||||| |||||||||| |||||||||||||||||||||
Db  435 CTACACAGTCAACCTTCTACAGCAGTGTACTTTGGGAACTTCCATGCAAGCCATGTCAT 494

Qy  530 CCTCGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGC 589
          ||||||||||||||||||||||||||||| ||||| ||||||||| ||||||||||||
Db  495 CCTCGCCATGTCCACGCCACCAACAGTGCCCTTCATCCTGGCCTTCATCAGTCTGGAAAAGGT 554

Qy  590 AGTCTATGTGGGCGTCTGGATCCCTGGATCCCCTGCCCCTCCGACTATACCTGACTTCATCTTTGC 649
          ||||| ||||||||| |||||||||||| ||||| |||||||||||| |||||||||||| |||||
Db  555 GGTCTATGTTGGCGTCATCCTGGTTCATGGGTCATGGGCATAGTCGGCATTGTGGG 614

Qy  650 CGAGCTCAGCCAGGGGACAATCAG 673
          ||||| |||||||||| |||||||
Db  615 CAACGTCAGTGAGGCAGATGACAG 638
```

RESULT 10
US-08-801-238-3
; Sequence 3, Application US/08801238
; Patent No. 5919896
GENERAL INFORMATION:
APPLICANT: Lee, James
APPLICANT: Wood, William I.
TITLE OF INVENTION: PF4A RECEPTOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,238
FILING DATE: 19-Feb-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284586
FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-801-238-3

Query Match 64.1%; Score 439.2; DB 2; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

```
Qy   50 GCAGGTAGCAGTGACCCCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTGTAGAGC 109
          |||| |||||||||| |||   ||| |||||||||||| ||||||  |||||||  ||  
Db   20 GCGGCGGCGCAAAGTGACGGCCGAGGGCCTGAGTGCTCCAGTAGCGTCCCAGTAGGAAAC  79

Qy  110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
          |||    ||||||||||     ||||  |||||||||||||| |||||||||  ||| ||
Db   80 CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134

Qy  170 AGTCGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
          | ||||||| ||||||||||||||||  |||||||||||| ||||| ||||||| ||  
Db  135 AATGGGCTCAGGGGACTATGACTCCATGAAGAAGAACCCTGTTTCCGTGAAGAAATGCTAA 194

Qy  230 TTTCAATAGGATCTTCCTGCCCACCATCACTTCTTTCATCATCATCTTCTTAACTGGCATAGTCGG 289
          ||||||| |||| ||||||||||||||||| |||||||||||| |||||| ||| |||||||||||
Db  195 TTTCAATAAAATCTTCCTGCCCACCATCACTTCTTTCATCATCATCTTCTTAACTGGCATTGTGGG 254

Qy  290 CAATGGATTGGTGATCCTGGTTACCAGAAGAAGCTAAGGAGCATAGGACGGACAA 349
          ||||| |||||| |||||| ||||||| |||||| ||||| |||| |||| |||
Db  255 CAATGGATTGGTCATCCTGGTTACCAGAAGAAACTGAGAAGCATGAGCATGACGACAA 314

Qy  350 GTACCGGCTGCCACCTGTCAGTGCTGGCTGACCCTCCTTGTCATCATCACACTCCCCTTCTGGGC 409
          ||||||| |||||| |||||| |||||  ||||| ||| |||||||||||  |||||| ||||||
Db  315 GTACAGGCTGCCACCTGTCAGTGGGCCGAGACTTGGTAAGGCTGTCCATATCAT 374

Qy  410 AGTTGATGCCATGGCCTGACTGGTACTTTGGGAAATTTTCTGTGTAAGGCTGTCCATATCAT 469
          |||||||||||||||| |||||||||||||||||||||||||| ||||| ||||||||||||
Db  375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTTCCATGTCAT 434

Qy  470 CTACACTGTCAACCTTCTACAGCAGCTTCATCCTGGCCTTCATCAGCCTGGACCCGGTA 529
          ||||||||||||||||| |||||||| |||||||||| |||||||||||||||||||||
Db  435 CTACACAGTCAACCTTCTACAGCAGTGTACTTTGGGAACTTCCATGCAAGCCATGTCAT 494

Qy  530 CCTCGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGC 589
          ||||||||||||||||||||||||||||| ||||| ||||||||| ||||||||||||
Db  495 CCTCGCCATGTCCACGCCACCAACAGTGCCCTTCATCCTGGCCTTCATCAGTCTGGAAAAGGT 554

Qy  590 AGTCTATGTGGGCGTCTGGATCCCTGGATCCCCTGCCCCTCCGACTATACCTGACTTCATCTTTGC 649
```

Db   135   AATGGGCTCAGGGGACTATGACTTCCATGAAGGAACCCTGTTTCCGTGAAGAAATGCTAA 194
            ||||||| |||||||||||||| ||||||||||||||||||||| || ||||||||||||

Qy   230   TTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTTCGG 289
            ||||||| |||||||||||||| |||||||||||||||||||||| |||||||||

Db   195   TTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTTCTTAACTGGCATTGTGGG 254

Qy   290   CAATGGATTGGTGATCCTGGTCATGGGTTACCAGGAGGAGCATGACGGACAA 349
            |||||||| |||||| ||||||||||||||||||||||||||||||||||||

Db   255   CAATGGATTGGTCATCCTGGTCATGGGTTACCAGGAAGAACTGAGAGCATGACGGACAA 314

Qy   350   GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTTCTGGGC 409
            |||||| ||||||||||||||||||||||||||||||| ||||||| ||||||||||||

Db   315   GTACAGGCTGCACCTGCCAGTGCGGCCGACCTCCTCTTGTCATCACGCTTCCCTTCTGGGC 374

Qy   410   AGTTGATGCGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGTGTCCATATCAT 469
            |||||||| |||||||||||| |||||||||||||| |||||||||||||||||||||

Db   375   AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy   470   CTACACTGTCAAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTA 529
            |||| ||||||| || |||||||||||||||||||||||||||||||||||||||||

Db   435   CTACACAGTCAAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTTGGACCGCTA 494

Qy   530   CCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTCGGCTGAAAAGGC 589
            ||| ||||||||||||||||| |||||| |||||||||||||||||||||||||||||

Db   495   CCTGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGT 554

Qy   590   AGTCTATGTGGGCGTCTGGATCCCAGCCCCTCCTCGACTATACCTGACTTCATCTTTGC 649
            ||| ||||| ||||||||||||||| |||||||||||| ||| |||||||||||||||

Db   555   GGTCTATGTTGGGCGTCTGGATCCCTGCCCCTCCTCGGATCCCTGCTATTCCCGACTATTCCCGACTTCATCTTTGC 614

Qy   650   CGACGTCAGCCAGGGGGACATCAG 673
            ||| ||| ||||||||||| ||

Db   615   CAACGTCAGTGAGGCCAGATGACAG 638

RESULT 12
US-09-104-296-3
; Sequence 3, Application US/09104296
; Patent No. 6087475
; GENERAL INFORMATION:
;- APPLICANT: Lee, James
;- APPLICANT: Wood, William I.
;- TITLE OF INVENTION: PF4A Receptors
;- NUMBER OF SEQUENCES: 6
;- CORRESPONDENCE ADDRESS:
;-- ADDRESSEE: Genentech, Inc.
;-- STREET: 1 DNA Way
;-- CITY: South San Francisco
;-- STATE: California
;-- COUNTRY: USA
;-- ZIP: 94080
;- COMPUTER READABLE FORM:
;-- MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;-- COMPUTER: IBM PC compatible
;-- OPERATING SYSTEM: PC-DOS/MS-DOS
;-- SOFTWARE: WinPatin (Genentech)
;- CURRENT APPLICATION DATA:
;-- APPLICATION NUMBER: US/09/104,296
;-- FILING DATE: 24-June-1998
;-- CLASSIFICATION:
;- PRIOR APPLICATION DATA:
;-- APPLICATION NUMBER: 08/701265
;-- FILING DATE: 22-AUG-1996
;- PRIOR APPLICATION DATA:
;-- APPLICATION NUMBER: 08/664228
;-- FILING DATE: 06-JUN-1996
;- PRIOR APPLICATION DATA:
;-- APPLICATION NUMBER: 08/076093
;-- FILING DATE: 11-JUN-1993
;- PRIOR APPLICATION DATA:
;-- APPLICATION NUMBER: 07/810782
;-- FILING DATE: 19-DEC-1991
;- ATTORNEY/AGENT INFORMATION:

Db   555   GGTCTATGTTGGGCGTCTGGATCCCTGCCCCTCCTCGGATCCCTGCTATTCCCGACTTCATCTTTGC 614

Qy   650   CGACGTCAGCCAGGGGGACATCAG 673
            ||| ||| |||||||||||||

Db   615   CAACGTCAGTGAGGCAGATGACAG 638

RESULT 11
US-08-801-228-3
; Sequence 3, Application US/08801228
; Patent No. 5922541
; GENERAL INFORMATION:
;- APPLICANT: Lee, James
;- APPLICANT: Wood, William I.
;- TITLE OF INVENTION: METHODS FOR DETECTION AND AMPLIFICATION OF
;- TITLE OF INVENTION: PF4A RECEPTOR NUCLEIC ACID
;- NUMBER OF SEQUENCES: 6
;- CORRSPONDENCE ADDRESS:
;-- ADDRESSEE: Genentech, Inc.
;-- STREET: 460 Point San Bruno Blvd
;-- CITY: South San Francisco
;-- STATE: California
;-- COUNTRY: USA
;-- ZIP: 94080
;- COMPUTER READABLE FORM:
;-- MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;-- COMPUTER: IBM PC compatible
;-- OPERATING SYSTEM: PC-DOS/MS-DOS
;-- SOFTWARE: WinPatin (Genentech)
;- CURRENT APPLICATION DATA:
;-- APPLICATION NUMBER: US/08/801,228
;-- FILING DATE: 19-Feb-1997
;-- CLASSIFICATION: 435
;- PRIOR APPLICATION DATA:
;-- APPLICATION NUMBER: 08/284586
;-- FILING DATE: 10-AUG-1994
;- PRIOR APPLICATION DATA:
;-- APPLICATION NUMBER: 08/076093
;-- FILING DATE: 11-JUN-1993
;- PRIOR APPLICATION DATA:
;-- APPLICATION NUMBER: 07/810782
;-- FILING DATE: 19-DEC-1991
;- ATTORNEY/AGENT INFORMATION:
;-- NAME: Love, Richard B.
;-- REGISTRATION NUMBER: 34,659
;-- REFERENCE/DOCKET NUMBER: P0706P2P1D3
;- TELECOMMUNICATION INFORMATION:
;-- TELEPHONE: 415/225-5530
;-- TELEFAX: 415/952-9881
;-- TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
;- SEQUENCE CHARACTERISTICS:
;-- LENGTH: 1737 base pairs
;-- TYPE: Nucleic Acid
;-- STRANDEDNESS: Single
;-- TOPOLOGY: Linear
US-08-801-228-3

Query Match   64.1%;   Score 439.2;   DB 2;   Length 1737;
Best Local Similarity   82.7%;   Pred. No. 1.4e-114;
Matches 516;   Conservative   0;   Mismatches 103;   Indels   5;   Gaps   1;

Qy   50   GCAGGTAGCAGTGACCCTCTGAGGCGTTGTGCTCCGGTAACCACCACGGCTGTAGAGC 109
            ||||||||||||| ||||| |||| ||||||||||||||||||||

Db   20   GCGCGGCGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTTAGCCACGCGCATCTGGAGAAC 79

Qy   110   GAGTGTTGCCATGGAACCGATCAGTCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
            ||||||| | |||||||||||||||||||||||||

Db   80   CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACACCGAGGA 134

Qy   170   AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCCGGGGATGAAAACGTCCA 229
            ||| || || |||||||||| |||||| ||||| |||||

```
;    TITLE OF INVENTION:    METHODS FOR TREATING INFLAMMATORY DISRODERS
;    NUMBER OF SEQUENCES: 3
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:   460 Point San Bruno Blvd
;       CITY:   South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/06380
;       FILING DATE:
;       CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/076093
;       FILING DATE:  11-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;       NAME:  Love, Richard B.
;       REGISTRATION NUMBER:  34,659
;       REFERENCE/DOCKET NUMBER:  706P2P1
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-5530
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;    INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  1737 bases
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
PCT-US94-06380-2
```

```
Query Match        64.1%;   Score 439.2;  DB 5;   Length 1737;
Best Local Similarity  82.7%;  Pred. No. 1.4e-114;
Matches  516;  Conservative   0;  Mismatches  103;  Indels   5;  Gaps   1;

Qy    50  GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC  109
          ||||||| |||||| ||| ||||||| ||||| ||| |||||||| |||||| ||| ||
Db    20  GCGCGGCGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCACCGGCTGGAGAAC   79

Qy   110  GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA  169
          || ||| |||||||| |||  |||||||  |||| ||||||||||||||||||||||||||
Db    80  CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA  134

Qy   170  AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA  229
          | |||  ||||||||||||||||||||||||||||||||||||| ||||||||||| |||
Db   135  AATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA   194

Qy   230  TTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTTGACTGGCATGGTCGG  289
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   195  TTTCAATAAAATCTTCCTGCCCACCATCACTTCTTCATCTTCTTAACTGGCATTGGTGGG  254

Qy   290  CAATGGATTGGTGATCCTGGTCATGGGTCATGGGTACCAGAAGAAGCTAAGGAGCATGAC  349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   255  CAATGGATTGGTCATCCTGGTCATGGGTACCAGAAGAAACTGAGAAGCATGAC      314

Qy   350  GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGC  409
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   315  GTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCACGCTCCCCTTCTGGGC  374

Qy   410  AGTTGATGCCATGGCTGACTGGTACTTTGGTAAGGCTGTCCATATCAT  469
          ||||||| ||||||| ||| |||||| ||||||||||| ||||||||
Db   375  AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGCAGTCCATGTCAT  434

Qy   470  CTACACTGTCAACCTCTACAGCAGCCTTCATCAGCCTCATCAGCCTTCATCAGCCTTCA  529
          |||||||||||||||||||||||||||||||||
Db   435  CTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTCAGTGTTGGACCTTCATCAGCCTTCA  494
```

```
RESULT 13
PCT-US94-06380-2
; Sequence 2, Application PC/TUS9406380
; GENERAL INFORMATION:
;    APPLICANT:  Chuntharapai, Anan
;    APPLICANT:  Lee, James
;    APPLICANT:  Hebert, Caroline
;    APPLICANT:  K. Jin Kim
;    APPLICANT:  Genentech, Inc.
```

```
NAME:  Love, Richard B.
REGISTRATION NUMBER:  34,659
REFERENCE/DOCKET NUMBER:  P0706P2C2
TELECOMMUNICATION INFORMATION:
   TELEPHONE:  415/225-5530
   TELEFAX:  415/952-9881
   TELEX:  910/371-7168
INFORMATION FOR SEQ ID NO:  3:
SEQUENCE CHARACTERISTICS:
   LENGTH:  1737 base pairs
   TYPE:  Nucleic Acid
   STRANDEDNESS:  Single
   TOPOLOGY:  Linear
US-09-104-296-3
```

```
Query Match        64.1%;   Score 439.2;  DB 3;   Length 1737;
Best Local Similarity  82.7%;  Pred. No. 1.4e-114;
Matches  516;  Conservative   0;  Mismatches  103;  Indels   5;  Gaps   1;

Qy    50  GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGC  109
          ||||||| |||||| ||| ||||||| ||||| ||| |||||||| |||||| ||| ||
Db    20  GCGCGGCGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCACCGGCTGGAGAAC   79

Qy   110  GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA  169
          || ||| |||||||| |||  |||||||  |||| ||||||||||||||||||||||||||
Db    80  CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA  134

Qy   170  AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA  229
          | |||  |||||||||||||||||||||||||||||||||||| ||||||||||| |||
Db   135  AATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA   194

Qy   230  TTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTTGACTGGCATGGTCGG  289
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   195  TTTCAATAAAATCTTCCTGCCCACCATCACTTCTTCATCTTCTTAACTGGCATTGGTGGG  254

Qy   290  CAATGGATTGGTGATCCTGGTCATGGGTCATGGGTACCAGAAGAAGCTAAGGAGCATGAC  349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   255  CAATGGATTGGTCATCCTGGTCATGGGTACCAGAAGAAACTGAGAAGCATGACGAGCAGGACAA  314

Qy   350  GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGC  409
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   315  GTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCACGCTCCCCTTCTGGGC  374

Qy   410  AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT  469
          ||||||| ||||||| ||| |||||| |||   ||| ||||| |||||| |||||||||
Db   375  AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT  434

Qy   470  CTACACTGTCAACCTCTACAGCAGCCTTCATCCTGGCCTTCATCAGCCTGGACCGGTA  529
          |||||||||||||||||||||||||| |||| || ||||| |||||| ||  |||
Db   435  CTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTCAGTGTTGGACCTTCATCAGCCTGGACCGCTA  494

Qy   530  CCTCGCCATTGTCACGCCACCAACAGTCAAGCCAAGGAAACTGCTGGCTGAAAAGGC  589
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   495  CCTGGCCATCGTCACGCCACCAACAGTCAGGCCAAGGAAGCTGTTGGCTGAAAAGGT  554

Qy   590  AGTCTATGTGGGGCGTCGGATCTGGATCTCCGATCCCGCCTCGCTGCTACATACCTTTGC  649
          ||||||| |||| ||||  ||| ||||||||||||| ||||||||| |||||| ||||||
Db   555  GGTCTATGTGGGGCGTCTGGATCTGGATCTCCCGACTATTCCCGACTATTCCCATCTTTGC  614

Qy   650  CGAGGTCAGCCAGGGGGACATCAG  673
          ||||| |||||||  |||| |||
Db   615  CAACGTCAGTGAGGCAGAGATGACAG  638
```

Qy 256 TCTACTTCATCATCTTCTTGACTGGCCATAGTCGGCAATAGTCGGCATGGATCCTGGTCATGG 315
       ||||||||| |||||||||| |||||| ||||| ||||||||| ||||| |||||||||||
Db 156 TCTACTTCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGG 215

Qy 316 GTTACCAGAAGAAGCTAAGGAGCATGAGCGGACGACAAGTACCGGCTGCACCTGTCAGTGGCTG 375
       ||||||||| || ||||| ||||||||| ||||| ||||||| ||||||||||||||||| |
Db 216 GTTACCAGAGAACTGAGAAGCATGAGCATGACGACAGTACCAGGGTGCACCTGTCAGTGGCCG 275

Qy 376 ACCTCCTCTTGTCATCACACTCCCCTTCTGGGCAGTGATGATGCCATGGCTGACGACTGGTACT 435
       |||||| |||||||||||| |||||||| ||||||||||||||||||||||| |||||||||
Db 276 ACCTCCTTTGTCATCACGCCTCCCTTCTCGGCAGTGATGCCGGTGATGCCGTGGCAAACTGGTACT 335

Qy 436 TTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTCACAGCAGCG 495
       |||||| | || ||||| |||||| |||| ||||||| || ||||||||||||||||
Db 336 TTGGGAACTTCCTATGCAAGGCAGTCCATTGCATCATCTACACAGTCAACCTCTCACAGCAGTG 395

Qy 496 TTCTCATCCTGGCCTTCATCAGCCTGACCACCTCCGCATTGTCCACGCCACCAACA 555
       |||||||||||||||| |||||| ||||||||| |||||||||||||||||| ||||||
Db 396 TCCTCATCCTGGCCTTCATCAGTCTGGCCACCTGTCCACCGCCATCGTCCACGCCACCAACA 455

Qy 556 GTCAAAGCCAAGGAAACTGCTGGCTGAAAAGCCAGTCTATGTGGGCGTCTCGGATCCCAG 615
       |||| |||||||||| ||||||||| ||||| ||||||| ||||||||||||||||||||
Db 456 GTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAAGGTGGTCTATGTTGGGCGTCTGGATCCCTG 515

Qy 616 CCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCAGCGGGGGACATCAG 673
       ||||| ||||||| || |||| |||||||||||||| |||||||||||||||||| |||
Db 516 CCCTCCTGCTGCTGCTATTCCCGACTTCATCTTTGCCAACGTCAGTGAGGCAGATGACAG 573

RESULT 15
US-08-153-848-45
; Sequence 45, Application US/08153848
; Patent No. 5759804
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;   ADDRESSEE: Bicknell
;   STREET: 6300 Sears Tower, 233 South Wacker Drive
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60606
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/153,848
;   FILING DATE:
;   CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/977,452
;   FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
;   NAME: No. 5759804and, Greta E.
;   REGISTRATION NUMBER: 35,302
;   REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (312) 474-6300
;   TELEFAX: (312) 474-0448
;   TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1317 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear

Qy 530 CCTCGCCATTGTCCACGCCACCAACAGTCAAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGC 589
       ||||||||||||| ||||||||||||||||| || ||||||||||| ||||||||||||||||
Db 495 CCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGT 554

Qy 590 AGTCTATGTGGGCGTCTCGGATCCCTCCTGACTTCATCTTTGC 649
       ||||||||| ||||||||||||||||||||| |||||||||||
Db 555 GGTCTATGTTGGGCGTCTGGATCCCTGCTGACTTCATCTTTGC 614

Qy 650 CGACGTCAGCCAGGGGGACATCAG 673
       |||||||||| ||||||||| |||
Db 615 CAACGTCAGTGAGGCAGATGACAG 638

RESULT 14
US-09-016-434-1235
; Sequence 1235, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/016,434
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1235:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1225 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g189313
US-09-016-434-1235

Query Match   62.9%;   Score 430.8;   DB 4;   Length 1225;
Best Local Similarity   87.5%;   Pred. No. 3e-112;
Matches   471; Conservative   0; Mismatches   67; Indels   0; Gaps   0;

Qy 136 TGAGTATATACACTTCTGATAACTACCTCTGAAGAAGTGGGGTCTGAGACTATGACTCCA 195
       |||||||||||||||||| |||||| |||| |||||||| |||| ||||||||||||||
Db 36 TCAGTATATACACTTCAGATAACTACCACCGGAAATGGGCTCAGGGGACTATGACTCCA 95

Qy 196 ACAAGGAACCCTGCTTCCGGATGAAAGTCCATTCATTCAATAGGATCTTCCTGCCCACCA 255
       || |||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db 96 TGAAGGAACCCTGTTTCCGGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCCACCA 155

Query Match   62.5%;   Score 428;   DB 1;   Length 1317;
Best Local Similarity   87.0%;   Pred. No. 1.9e-111;
Matches 470;   Conservative 0;   Mismatches 70;   Indels 0;   Gaps 0;

```
Qy    134 TGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTCGAGAGACTATGACTC 193
          |||  ||||||||||||| |||||||||   ||||||| |||||| |||||||||||||||
Db    164 TTTGCAGATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTC 223

Qy    194 CAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCAC 253
          || |||||||||||| ||||| ||||||||||| ||| ||||||| |||||||||||||
Db    224 CATGAAGGAACCCTGTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCCAC 283

Qy    254 CATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCCAATGGATTGGTGATCCTGGTCAT 313
          |||||||||||||||||||||| |||||||| ||| ||||||||||||| || |||||||
Db    284 CATCTACTTCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTCAT 343

Qy    314 GGGTTACCAGAAGAAGCTAAGGAGGCAATGACGGACAAGTACCGGCTGCACCTGTCAGTGGC 373
          ||| |||||||||||||| ||  |  | ||||||||||||||||||||||||||||||||||
Db    344 GGGGTTACCAGGAAGAAGCTGAGAAACTGAGAAGCAATGACGGACAAGTACCGGCTGCACCTGTCAGTGGC 403

Qy    374 TGACCTCCTCTTTGTCATCACACTCCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTA 433
          ||||||||||||||||||||| |||||||| || |||||||||||||| ||||| ||||||
Db    404 CGACCTCCTCTTTGTCATCACGCTCCCCTTCTCCGTGGGCAGTTGATGCCGTGGCAAACTGGTA 463

Qy    434 CTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCCACACTTGTCAACCTCTACAGCAG 493
          ||||||||| || ||| || ||||| |||| ||||| || ||| |||||||||||||||||
Db    464 CTTTGGGAACTTCCTATGCAAGGCAGTCATGTCATCTCATCTACACAGTCAACCTCTACAGCAG 523

Qy    494 CGTTCTCATCCTGGCCTTCATCAGCCTGGACCGGTCACCTCCGCCATTGTCCACGCCACCAA 553
           || ||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||
Db    524 TGTCCTCATCCTGGCCTTCATCAGCCTGGACCGGTCACCTCCGGCCATCGTCCACGCCACCAA 583

Qy    554 CAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAAGGCAGTCTATGTGGGGCGTCTGGATCCC 613
          ||||||| ||||||||| || |  ||||||||||| | |||||||| |||||||||||||||
Db    584 CAGTCAGAGGCCAAGGAAGCTGTTCGGCTGAAAAAGGTGGTCTATGTTGGGCGTCTGGATCCC 643

Qy    614 AGCCCTCCTCCTGACTATACCTGACTCATCTTTGCCGACGTCAGCAGGGGACATCAG 673
           |||| ||| |||||||| || ||| |||||||||| ||||||| |||| |||| |||
Db    644 TGCCCCCCTGCTGACTATTCCCGACTTCATCTTTGCCAACGTCAGTGAGGCAGATGACAG 703
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 17, 2004, 12:11:00 ; Search time 335.238 Seconds
                                        (without alignments)
                                        9271.041 Million cell updates/sec

Title: US-09-367-052-5
Perfect score: 685
Sequence: 1 ccatcctaatacgactcact.......gacatccagtcaggggatga 685

Scoring table: IDENTITY NUC
               Gapop 10.0 , Gapext 1.0

Searched: 294734 seqs, 226924515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published_Applications_NA:*
    1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
    2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
    3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
    4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
    5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
    6: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
    7: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
    8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
    9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
   10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
   11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
   12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
   13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
   14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
   15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
   16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
   17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
   18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
   19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	446.8	65.2	1679	13	US-10-211-462-80	Sequence 80, Appl
2	446.8	65.2	1679	13	US-10-188-906-9	Sequence 9, Appli
3	446.8	65.2	1679	13	US-10-342-887-912	Sequence 912, App
4	446.8	65.2	1679	13	US-10-151-274-14	Sequence 14, Appl
5	446.8	65.2	1679	13	US-10-172-118-912	Sequence 912, App
6	446.8	65.2	1679	13	US-10-170-385-332	Sequence 332, App
7	446.8	65.2	1679	15	US-10-225-567A-75	Sequence 75, Appl
8	446.8	65.2	1679	15	US-10-021-660-58	Sequence 58, Appl
9	446.8	65.2	1679	16	US-10-341-434-177	Sequence 177, App
10	446.8	65.2	1711	10	US-09-971-392-20	Sequence 20, Appl
11	446.8	65.2	1711	15	US-10-641-510-459	Sequence 453, App
12	445.8	65.1	1664	17	US-10-641-643-1213	Sequence 1213, Ap
13	439.2	64.1	1737	9	US-09-104-063-3	Sequence 3, Appli
14	439.2	64.1	1737	13	US-10-666-689-3	Sequence 3, Appli
15	437.6	63.9	1670	9	US-09-880-107-2143	Sequence 2143, Ap
16	437.6	63.9	1670	10	US-09-960-706-636	Sequence 636, App
17	437.6	63.9	1670	16	US-10-372-683-1	Sequence 1, Appli
18	437.6	63.9	1670	16	US-10-440-464-77	Sequence 77, Appl
19	432.4	63.1	1102	9	US-09-870-759-143	Sequence 143, App
20	432.4	63.1	1102	10	US-09-751-708A-143	Sequence 143, App
21	430.8	62.9	1059	11	US-09-826-509-484	Sequence 484, Appl
22	430.8	62.9	1059	13	US-09-813-651B-84	Sequence 84, Appl
23	430.8	62.9	1059	15	US-10-014-322A-125	Sequence 125, App
24	430.8	62.9	1059	16	US-10-160-401-2	Sequence 2, Appli
25	430.8	62.9	1225	15	US-10-101-510-741	Sequence 741, App
26	430.8	62.7	1225	16	US-10-305-720-1235	Sequence 1235, Ap
27	429.2	62.7	5161	16	US-10-160-401-1	Sequence 1, Appli
28	418.8	61.1	1224	15	US-10-101-510-674	Sequence 674, App
29	318.8	46.5	1076	15	US-10-017-161-789	Sequence 789, App
30	318.8	46.5	1076	16	US-10-292-798-687	Sequence 687, App
31	246	35.9	1902	9	US-09-953-692-1	Sequence 1, Appli
32	246	35.9	1902	9	US-09-953-717-1	Sequence 1, Appli
33	152.4	22.2	1119	12	US-09-997-722-162	Sequence 162, App
34	152.4	22.2	1119	15	US-10-251-385-65	Sequence 65, Appl
35	152.4	22.2	1119	15	US-10-251-385-199	Sequence 199, App
36	152.4	22.2	1679	9	US-09-104-063-5	Sequence 5, Appli
37	152.4	22.2	1679	13	US-10-666-689-5	Sequence 5, Appli
38	152.4	22.2	2824	12	US-09-997-722-161	Sequence 161, App
39	152.4	22.2	2824	15	US-10-225-567A-59	Sequence 59, Appl
40	152.4	22.2	3011	15	US-10-240-965-154	Sequence 154, App
41	152.4	22.2	3620	13	US-09-997-774-893	Sequence 893, App
42	152.4	22.2	32404	12	US-09-997-722-160	Sequence 160, App
43	141.6	20.7	172	9	US-09-796-692-2641	Sequence 2641, Ap
44	141.6	20.7	172	15	US-10-040-862-2641	Sequence 2641, Ap
45	141.6	20.7	172	16	US-10-057-475B-2641	Sequence 2641, Ap

ALIGNMENTS

RESULT 1
US-10-211-462-80
; Sequence 80, Application US/1021462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-211-462-80

Query Match   65.2%; Score 446.8; DB 13; Length 1679;
Best Local Similarity   83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative   0; Mismatches 102; Indels   5; Gaps   1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTG 103
              ||| ||||||||| ||||| ||| |||||||||| ||| ||||||||| ||| |||||||

Db 12 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy 104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
Db 72 GAGAACCAGCGGTTACCATGGA----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy 164 TGAAGAAGTGGGGTCTGGAGACTATGACTATGACTTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
Db 127 CGAGGAAATGGGCTCAGGGGACTATGACTTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy 224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
Db 187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTTCTTAACTGGCAT 246

Qy 284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC 343
Db 247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGCATGAC 306

Qy 344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTT 403
Db 307 GGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCCTTCCCTT 366

Qy 404 CTGGGCAGTTGATGCCATGGCTGACTTGGTACTTTGGTAAGGCTGTCCA 463
Db 367 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCATGCAAGGCAGTCCA 426

Qy 464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
Db 427 TGTCATCTCACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy 524 CCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
Db 487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy 584 AAAGGCAGTCTATGTGGGCGTCTGGACCCCAGCCCTCCTCCGACTATACCTGACTTCAT 643
Db 547 AAAGGTGGTCTATGTTGGCGTCTGGATCCCCTGCTGCTGACTATTCCCGACTTCAT 606

Qy 644 CTTTGCCGACGTCAGCCAGGGGACCATCAG 673
Db 607 CTTTGCCAACGTCAGTGAGGCAGATGACAG 636

RESULT 2
US-10-181-906-9
; Sequence 9, Application US/10181906
; Publication No. US2004053864A1
; GENERAL INFORMATION:
; APPLICANT: Karsenty, Gerard
; APPLICANT: Amling, Michael
; APPLICANT: Ducy, Patricia
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BONE FORMATION VIA
; TITLE OF INVENTION: MODULATION OF NEUROPEPTIDE Y ACTIVITY
; FILE REFERENCE: 9142-020-999
; CURRENT APPLICATION NUMBER: US/10/181,906
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/02040
; PRIOR FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-181-906-9

Query Match   65.2%;   Score 446.8;   DB 13;   Length 1679;
Best Local Similarity   83.0%;   Pred. No. 2.5e-132;
Matches   523;   Conservative   0;   Mismatches 102;   Indels   5;   Gaps   1;

Qy 44 GCAGGTGCAGGTAGCAGTGACGTGACCCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACGGGCTG 103
Db 12 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Db 12 ...

Qy 104 ...

RESULT 3
US-10-342-887-912
; Sequence 912, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-912

Query Match   65.2%;   Score 446.8;   DB 13;   Length 1679;

```
Query Match         65.2%;  Score 446.8;  DB 13;  Length 1679;
Best Local Similarity  83.0%;  Pred. No. 2.5e-132;
Matches  523;  Conservative  0;  Mismatches 102;  Indels  5;  Gaps  1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
          ||||| ||||||||  |  |  ||||||   | ||  ||| |||
Db   12 GCGCGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy  104 TAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
          || |||| ||  ||||||||       |||||  ||||||||| ||||||| ||||||| 
Db   72 GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy  164 TGAAGAAGTGGGGTCTGGAGACTATGACTCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
          |||| ||| ||| |||| |||||||||||  |  ||||||||| |||||| |||||||
Db  127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy  224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCTTCTTGACTGCTGGCAT 283
          ||| |||||||||| ||||||||||||||||||  |||||||||| ||||| |||||
Db  187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCCTCATCTTCTTAACTGGCAT 245

Qy  284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC 343
          | || ||||||||||||| |||||||||||||||||||||||||| || |||||||||||
Db  247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAGCATGAC 306

Qy  344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCCTCTTTGTCATCACACTCCCCTT 403
          |||||||||| ||||||||||||||||  | ||||||| ||||||||||  |||| |||
Db  307 GGACAAGTACAGGCTGCACCTGTCAGTGCCGACCCTCCTTCTTTGTCATCACGCTTCCCTT 366

Qy  404 CTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGGAAATTTTGTGTAAGGCTGTCCA 463
          |||||||||||||||   |||| || |||||||||||||   |||  ||||||| |||||
Db  367 CTGGGCAGTTGATGCCGGTGGCAAACTGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy  464 TATCATCTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
          | ||||||||| |||||||||||||||||   | ||||||||||||||||||||| |||||
Db  427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy  524 CCGGTACCTGCCATTGTCCACAGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
          ||| ||||||| || |||||| |||||||||||||| |||||||||| |||| ||||||
Db  487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy  584 AAAGGCAGTCTATGTGGGCGTCTGGGCCTCCTGATCCCAGCCCCTCCTGACTATACCTCAT 643
          ||||| ||||||||| ||||||||| ||| |||| ||| ||||||||| |||||| | ||||
Db  547 AAAGGTGGTCTATGTTGGGCGTCTGGATCCCTGCCCCTCCTGCTGACTATTCCCGACTTCAT 606

Qy  644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673
          ||||||| |||||||| |||| |||| ||
Db  607 CTTTGCCAACGTCAGTGAGGCAGATGACAG 636
```

RESULT 4
US-10-151-274-14
; Sequence 14, Application US/10151274
; Publication No. US20030064071A1
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10/151,274
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/517,605
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-151-274-14

```
Best Local Similarity  83.0%;  Pred. No. 2.5e-132;
Matches  523;  Conservative  0;  Mismatches 102;  Indels  5;  Gaps  1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
          ||||| ||||||||  |  |  ||||||   | ||  ||| |||
Db   12 GCGCGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy  104 TAGAGCGGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
          || |||| ||  ||||||||       |||||  ||||||||| ||||||| ||||||| 
Db   72 GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy  164 TGAAGAAGTGGGGTCTGGAGACTATGACTCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
          |||| ||| ||| |||| |||||||||||  |  ||||||||| |||||| |||||||
Db  127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy  224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCTTCTTGACTGCTGGCAT 283
          ||| |||||||||| ||||||||||||||||||  |||||||||| ||||| |||||
Db  187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCCTCATCTTCTTAACTGGCAT 246

Qy  284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGCATGAC 343
          | || ||||||||||||| |||||||||||||||||||||||||| || |||||||||||
Db  247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAACTGAGAGCATGAC 306

Qy  344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCCTCTTTGTCATCACACTCCCCTT 403
          |||||||||| ||||||||||||||||  | ||||||| ||||||||||  |||| |||
Db  307 GGACAAGTACAGGCTGCACCTGTCAGTGCCGACCCTCCTTCTTTGTCATCACGCTTCCCTT 366

Qy  404 CTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGGAAATTTTGTGTAAGGCTGTCCA 463
          |||||||||||||||   |||| || |||||||||||||   |||  ||||||| |||||
Db  367 CTGGGCAGTTGATGCCGGTGGCAAACTGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy  464 TATCATCTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
          | ||||||||| |||||||||||||||||   | ||||||||||||||||||||| |||||
Db  427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGA 486

Qy  524 CCGGTACCTGCCATTGTCCACAGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
          ||| ||||||| || |||||| |||||||||||||| |||||||||| |||| ||||||
Db  487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy  584 AAAGGCAGTCTATGTGGGCGTCTGGGCCTCCTGATCCCAGCCCCTCCTGACTATACCTCAT 643
          ||||| ||||||||| ||||||||| ||| |||| ||| ||||||||| |||||| | ||||
Db  547 AAAGGTGGTCTATGTTGGGCGTCTGGATCCCTGCCCCTCCTGCTGACTATTCCCGACTTCAT 606

Qy  644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673
          ||||||| |||||||| |||| |||| ||
Db  607 CTTTGCCAACGTCAGTGAGGCAGATGACAG 636
```

RESULT 5
US-10-172-118-912
; Sequence 912, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens

Left column

Query Match   65.2%;   Score 446.8;   DB 13;   Length 1679;
Best Local Similarity   83.0%;   Pred. No. 2.5e-132;
Matches 523;   Conservative   0;   Mismatches 102;   Indels   5;   Gaps   1;

```
Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCGGTTCGGTAACCACCACGGCTG         103
        |||||  |||||||||||  | |  |||  |||||  ||||  ||||| || ||
Db   12 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGCCCAGTAGCCACCGCATCTG      71

Qy  104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
        |||| |  || |||| ||||||            ||| |||||||||||| ||||||| ||
Db   72 GAGAACCAGCCGGTTACCATGGA----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy  164 TGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAA    223
        ||| ||| ||||| ||  |||||||||||  |  ||||||||| | |||| |||||||
Db  127 CGAGGAAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAA    186

Qy  224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTGACTGGCAT   283
        |||| |||||||||  ||||||||||||||||||||  ||||||||||   ||||||||
Db  187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCACTCCATCATCATCTTAACTGGCAT   246

Qy  284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGTTACCAGAGAAGCTAAGGAGCATGAC   343
        |||| |||||||||||| |||||||||||||||||||||||||| | || ||||||||
Db  247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGTTACCAGAGAAACTGAGAAGCATGAC   306

Qy  344 GGACAAGTACCGGCTGACCTGTCAGTGCTGACCTCCTCTTGTCATCACACTCCCCTT    403
        ||||||||||  |||| |||||||||||||||||||||||||||||||| || ||||
Db  307 GGACAAGTACAGGCTGCACCTGTCAGTGCTGACCTCCTCTTGTCATCACGCTTCCCTT    366

Qy  404 CTGGGCAGTTGATGCCATGGCCTTGGGAAATTTTGTGTAAGGCTGTCCA          463
        ||||||||||||  ||  ||||  ||||||||| |||| ||||||| |||
Db  367 CTGGGCAGTTGACGCCGTGGCAAACTGGTTACTTTGGGAACTTCCATCGGCCTTCCA    426

Qy  464 TATCATCTACACTGTCAACCTCTACAGGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 523
        ||||||||||||||||| ||||||||| |||| ||||||||||||||||||||||| |||||||
Db  427 TGTCATCTACACAGTCAACCTCTACAGCAGTCAACCTCTACAGCAGTGTCCTCATCAGTCTGGA 486

Qy  524 CCGGTACCTGCCATTGTCCACCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGGA    583
        |||||||| |||| ||||||||||||||| |||||||||||||| |||| |||||||
Db  487 CCGCTACCTGGCCATGGCCATGGCCACCAACAAGGCCAAGGAAACTGCTGGCTGA    546

Qy  584 AAAGGCAGTCTATGTGGGGCGTCTGGATCCCAGCCCTCCTCCTCGACTATACCTGACTTCAT 643
        |||||||||||||| |||| |||  ||||| |||  ||||||| |||||| ||||||||||
Db  547 CCGCTACCTGGCCGTCTATGTTGGCCGTCTGGATCCCAGCCCTCCTCCTCGACTATTCCCGACTTCAT 606

Qy  644 CTTTGCCGACGTCAGCCAGGGGACATCAG                             673
        |||||||||||||||||||||||  | ||
Db  607 CTTTGCCAACGTCAGTCAGTGAGGCAGATGACAG                        636
```

; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David

Right column

Query Match   65.2%;   Score 446.8;   DB 13;   Length 1679;
Best Local Similarity   83.0%;   Pred. No. 2.5e-132;
Matches 523;   Conservative   0;   Mismatches 102;   Indels   5;   Gaps   1;

```
Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCGGTTGGTGTCCTGGTAACCACCACGGCTG   103
        |||||  |||||||||||  | |  |||  ||||  || | || ||||  ||||| || ||
Db   12 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGCCCAGTAGCCACCGCATCTG       71

Qy  104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
        |||| |  || |||| ||||||            ||| |||||||||||| ||||||| ||
Db   72 GAGAACCAGCCGGTTACCATGGA----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy  164 TGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGATGAAAA    223
        ||| ||| ||||| ||  |||||||||||  |  ||||||||| | |||| |||||||
Db  127 CGAGGAAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAA    186

Qy  224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCACTTCATCATCTTCTGACTGGCAT   283
        |||| |||||||||  ||||||||||||||||||||  ||||||||||   ||||||||
Db  187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCACTCCATCATCATCTTAACTGGCAT   246

Qy  284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGTTACCAGAGAAGCTAAGGAGCATGAC   343
        |||| |||||||||||| |||||||||||||||||||||||||| | || ||||||||
Db  247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGTTACCAGAGAAACTGAGAAGCATGAC   306

Qy  344 GGACAAGTACCGGCTGACCTGTCAGTGCTGACCTCCTCTTGTCATCACACTCCCCTT    403
        ||||||||||  |||| |||||||||||||||||||||||||||||||| || ||||
Db  307 GGACAAGTACAGGCTGCACCTGTCAGTGCTGACCTCCTCTTGTCATCACGCTTCCCTT    366

Qy  404 CTGGGCAGTTGATGCCATGGCCTTGGGAAATTTTGTGTCATCACACTCCCCTT       463
        |||||||||||| ||  ||||  ||||||||| |||||  || ||||
Db  367 CTGGGCAGTTGACGCCGTGGCAAACTGGTTACTTTGGGAACTTCCATCCTGGCCTTCCTT 426

Qy  464 TATCATCTACACTGTCAACCTCTACAGGCCACCAACAGTCTGCTGGA             523
        ||||||||||||||||| ||||||||| ||| |||||||
Db  427 TGTCATCTACACAGTCAACCTCTACAGCAGTCAACCTCTACAGCAGTCAGTCTGGA      486

Qy  524 CCGGTACCTGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGGA   583
        |||||||| |||| ||  ||||||||||||| ||||||||||||||||||||| |||||
Db  487 CCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAAAGGCCAAGGAAGAAGTCTGGA    546

Qy  584 AAAGGCAGTTCTATGTGGGGCGTCTGGATCCCAGCCCTCCTCCTCGACTATACCTGACATCAG 643
        ||||| |||||||| ||||| |||  ||||| |||  ||||||| |||||| ||||  |||||
Db  547 AAAGGTGTCTATGTTGGCCGTCTGGATCCCAGCCCTCCTCCTCGACTATTCCCGACTTCAT 606

Qy  644 CTTTGCCGACGTCAGCCAGGGGGGACATCAG                           673
        |||||||  |||||||| ||||| | || ||
Db  607 CTTTGCCAACGTCAGTCAGTGAGGCAGATGACAG                        636
```

; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.

; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-75

Query Match   65.2%;   Score 446.8; DB 15; Length 1679;
Best Local Similarity   83.0%;   Pred. No. 2.5e-132;
Matches   523; Conservative   0; Mismatches 102; Indels   5; Gaps   1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCCGTTGGTGCTCCGGTAACCACCACGGCTG 103
           |||| ||||||||| || ||||| |||||||||||||||||||| | |||||||||||||
Db   12 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy   104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAATACTC 163
           || |||||| ||||||||||||||||||||||| ||||||||||||||||| | || |
Db   72 GAGAACCAGCAGGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
           ||||||| ||||||||| |||||||||||||||| ||||||||| ||||| | |||||||
Db   127 CGAGGAAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy   224 CGTCCATTTCAATAGGATCTTCCTGCCCACCACCATCTACTTCATCATCATCTCTTGACTGGCAT 283
           ||||||||||||| |||||||||||||||||||||||||||||||||||||| | ||||||||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGGATTGGTGATCCTGGTTACCGGAGAAGCTAAGGAGCATGAC 343
           ||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCGGAGAAGAAACTGAGAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTT 403
           |||||||||||||||||||||||||| || ||||| |||||||||| ||||||||||||
Db   307 GGACAAGTACAGGCTGCACCTGTCAGTGGCGCCGACCTCCTCTTTGTCATCACGCTTCCCTT 366

Qy   404 CTGGGCAGTTGATGCCATGGCCATGGTACTTGGGAAATTTTTGTGTAAGGCTGTCCA 463
           ||||||||| ||||||||||||| ||||||||||||| ||||||||||||||||||
Db   367 CTGGGCAGTTGATGCCGTGGCAAACTTGGGAACTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCAATCCTGGCCTTCATCAGCCTGGA 523
           | |||||||||| |||||||||||||||||||| |||||||||| | |||||||||||||
Db   427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGA 486

Qy   524 CCGGTACCTCGCCATTGTCCACGCCACCAAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
           ||||||||||||||||||||||||||||||||||| |||||||| ||||||||| |||||
Db   487 CCGCTACCTGGCCATCGTCCACGCCACCAAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy   584 AAAGGCAGTCTATGTGGGCGTCTGGGATCCCAGCCCTCCTGACTATACCTGACTTCAT 643
           |||||| |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   547 AAAGGTGGTCTATGTTGGCGTCTGGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCAT 606

Qy   644 CTTTGCCGACGTCAGCAGCGGGGACATCAG 673
           |||||||||||||| ||| |||||||| |||
Db   607 CTTTGCCAACGTCAGTGGCGAGGCAGATGACAG 636

RESULT 8
US-10-021-660-58
; Sequence 58, Application US/10021660
; Publication No. US2003015292G6A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US

; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-021-660-58

Query Match   65.2%;   Score 446.8; DB 15; Length 1679;
Best Local Similarity   83.0%;   Pred. No. 2.5e-132;
Matches   523; Conservative   0; Mismatches 102; Indels   5; Gaps   1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTCTGAGCCGTTGGTGCTCCGGTAACCACCACGGCTG 103
           |||| ||||||||| || ||||| |||||||||||||||||||| | |||||||||||||
Db   12 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy   104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAATACTC 163
           || |||||| ||||||||||||||||||||||| ||||||||||||||||| | || |
Db   72 GAGAACCAGCAGGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy   164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
           ||||||| ||||||||| |||||||||||||||| ||||||||| ||||| | |||||||
Db   127 CGAGGAAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy   224 CGTCCATTTCAATAGGATCTTCCTGCCCACCACCATCTACTTCATCATCATCTCTTGACTGGCAT 283
           ||||||||||||| |||||||||||||||||||||||||||||||||||||| | ||||||||
Db   187 TGCTAATTTCAATAAAATCTTCCTGCCCACCATCTACTTCATCATCTCTTAACTGGCAT 246

Qy   284 AGTCGGCAATGGATTGGTGATCCTGGTTACCGGAGAAGCTAAGGAGCATGAC 343
           ||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCGGAGAAGAAACTGAGAGCATGAC 306

Qy   344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTT 403
           |||||||||||||||||||||||||| || ||||| |||||||||| ||||||||||||
Db   307 GGACAAGTACAGGCTGCACCTGTCAGTGGCGCCGACCTCCTCTTTGTCATCACGCTTCCCTT 366

Qy   404 CTGGGCAGTTGATGCCATGGCCATGGTACTTGGGAAATTTTTGTGTAAGGCTGTCCA 463
           ||||||||| ||||||||||||| ||||||||||||| ||||||||||||||||||
Db   367 CTGGGCAGTTGATGCCGTGGCAAACTTGGGAACTTGGGAACTTCCTATGCAAGGCAGTCCA 426

Qy   464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCAATCCTGGCCTTCATCAGCCTGGA 523
           | |||||||||| |||||||||||||||||||| |||||||||| | |||||||||||||
Db   427 TGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGA 486

Qy   524 CCGGTACCTCGCCATTGTCCACGCCACCAAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
           ||||||||||||||||||||||||||||||||||| |||||||| ||||||||| |||||
Db   487 CCGCTACCTGGCCATCGTCCACGCCACCAAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA 546

Qy   584 AAAGGCAGTCTATGTGGGCGTCTGGGATCCCAGCCCTCCTGACTATACCTGACTTCAT 643
           |||||| |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   547 AAAGGTGGTCTATGTTGGCGTCTGGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCAT 606

Qy   644 CTTTGCCGACGTCAGCAGCGGGGACATCAG 673
           |||||||||||||| ||| |||||||| |||
Db   607 CTTTGCCAACGTCAGTGGCGAGGCAGATGACAG 636

RESULT 9
US-10-341-434-177
; Sequence 177, Application US/10341434
; Publication No. US2003021583SA1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18

```
Query Match          65.2%;     Score 446.8;  DB 10;   Length 1711;
Best Local Similarity 83.0%;  Pred. No. 2.5e-132;
Matches 523;  Conservative   0;  Mismatches 102;  Indels   5;  Gaps   1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTG 103
         ||||| || ||||||| | ||||  ||||| ||| ||| |||| ||| |||| || |||
Db   47 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 106

Qy  104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGAGTATATACACTTCTGATAACTACTC 163
         || | | || || ||||   ||||  || |||| |||||||||||| |||||||||| |
Db  107 GAGAACCAGCGGTTACCATGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 161

Qy  164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
         | || || ||| || ||| |||||||||||| | |||||||| || ||||| || |||||
Db  162 CGAGGAAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 221

Qy  224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
         | | ||||||||||   ||||||||||||||||  |||  || |||||||||| ||||||
Db  222 TGCTAATTTCAATAAAAATCTTCCTGCCCACCATCCATCCATCTTCTTAACTGGCAT 281

Qy  284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGCTAAGGAGCATGAC 343
         | | ||||||||||||| |||||||||||||||| | |||||| | |||| ||||||
Db  282 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTCATCCAGAAGAAACTGAGAAGCATGAC 341

Qy  344 GGACAAGTACCGGCTGCACCTGGACTTGGTGTCATCACACTCCCTT 403
         ||||||||| || | || ||||  ||| | |||||||  |||||||
Db  342 GGACAAGTACAGGCCTGCACCTGTCAGTGCCGACCTCCTCTTTGTCATCACGCTTCCTT 401

Qy  404 CTGGGCAGTTGATGCCATGCGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCA 463
         ||||||||||||||| |||||| |  |||||||| |||||  ||  |  |||||| |||||
Db  402 CTGGGCAGTTGATGCCGTGCGCAAACTGGTACTTTTGGGAACTTCCTCATGCAAGGCAGTCCA 461

Qy  464 TATCATCTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
         | ||||| ||| ||||||||||||||  ||  |  ||||||||||||||||||| |||||
Db  462 TGTCATCTCACACAGTCAACCTCTACACGCGTCCCATCCTGGCCTTCATCAGTCTGGA 521

Qy  524 CCGGTACCTCGCCATTGTCAACCTCCTACAGCGTTCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
         ||| ||||||||||| ||| ||||| ||||| |||||||||||||||||||||||| |||||
Db  522 CCGCTACCTCGCCATCGTCCACCTCCACGCGCCACCAACAGTCAGACGAGGCCAAGGAAGCTGA 581

Qy  584 AAAGGCAGTCTATGTGGGCGTCTGGGCGTCTGGGGGGGGACATTCAT 643
         ||||| |||||||| ||| |||||||||||||||||| ||| ||||
Db  582 AAAGGTGGTCTATGTTGGGGCGTCTGGGCGTCTGGGGGGGACTTCAT 641

Qy  644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673
         ||||||| ||||||| || || | ||| |
Db  642 CTTTGCCAACGTCAGTGAGGCAGATGACAG 671
```

```
Query Match          65.2%;     Score 446.8;  DB 16;   Length 1679;
Best Local Similarity 83.0%;  Pred. No. 2.5e-132;
Matches 523;  Conservative   0;  Mismatches 102;  Indels   5;  Gaps   1;

Qy   44 GCAGGTGCAGGTAGCAGTGACCCTGAGGCGTTTGGTGCTCCGGTAACCACCACGGCTG 103
         ||||| || ||||||| | ||||  ||||| ||| ||| |||| ||| |||| || |||
Db   12 GCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 71

Qy  104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGAGTATATACACTTCTGATAACTACTC 163
         || | | || || ||||   ||||  || |||| |||||||||||| |||||||||| |
Db   72 GAGAACCAGCGGTTACCATGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 126

Qy  164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
         | || || ||| || ||| |||||||||||| | |||||||| || ||||| || |||||
Db  127 CGAGGAAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAA 186

Qy  224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
         | | ||||||||||   ||||||||||||||||  |||  || |||||||||| ||||||
Db  187 TGCTAATTTCAATAAAAATCTTCCTGCCCACCATCCATCCATCTTCTTAACTGGCAT 246

Qy  284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGCTAAGGAGCATGAC 343
         | | ||||||||||||| |||||||||||||||| | |||||| | |||| ||||||
Db  247 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTCATCCAGAAGAAACTGAGAAGCATGAC 306

Qy  344 GGACAAGTACCGGCTGCACCTGGACTTGGTGTCATCACACTCCCTT 403
         ||||||||| || | || ||||  ||| | |||||||  |||||||
Db  307 GGACAAGTACAGGCCTGCACCTGTCAGTGCCGACCTCCTCTTTGTCATCACGCTTCCTT 366

Qy  404 CTGGGCAGTTGATGCCATGCGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCA 463
         ||||||||||||||| |||||| |  |||||||| |||||  ||  |  |||||| |||||
Db  367 CTGGGCAGTTGATGCCGTGCGCAAACTGGTACTTTTGGGAACTTCCTCATGCAAGGCAGTCCA 426

Qy  464 TATCATCTACACTGTCAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
         | ||||| ||| ||||||||||||||  ||  |  ||||||||||||||||||| |||||
Db  427 TGTCATCTCACACAGTCAACCTCTACACGCGTCCCATCCTGGCCTTCATCAGTCTGGA 486

Qy  524 CCGGTACCTCGCCATTGTCAACCTCCTACAGCGTTCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGA 583
         ||| ||||||||||| ||| ||||| ||||| |||||||||||||||||||||||| |||||
Db  487 CCGCTACCTCGCCATCGTCCACCTCCACGCGCCACCAACAGTCAGACGAGGCCAAGGAAGCTGA 546

Qy  584 AAAGGCAGTCTATGTGGGCGTCTGGGCGTCTGGGGGGGGACATTCAT 643
         ||||| |||||||| ||| |||||||||||||||||| ||| ||||
Db  547 AAAGGTGGTCTATGTTGGGGCGTCTGGGCGTCTGGGGGGGACTTCAT 606

Qy  644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673
         ||||||| ||||||| || || | ||| |
Db  607 CTTTGCCAACGTCAGTGAGGCAGATGACAG 636
```

; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 459
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1706)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-459

Query Match 65.2%; Score 446.8; DB 15; Length 1711;
Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

```
Qy    44 GCAGGTGCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
            |||||||  |||||||||||||||  |||  || ||||||||||||  ||||||||||
Db    47 GCGGCAGCAGGTAGCAAAGTGACAGTGACCGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCCACCATCTG 106

Qy   104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCGATAACTACTC 163
            |||||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db   107 GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 161

Qy   164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAA 223
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   162 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGAACCCTGTTTCCGTGAAGAAAA 221

Qy   224 CGTCCATTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   222 TGCTAATTTCAATAAAATCTTCCTGCCACCATCTACTTCCATCATCTTCTTAACTGGCAT 281

Qy   284 AGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGTAAGGAGCATGAC 343
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   282 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTACCAGAAGAAGAAACTGAGAAGCATGAC 341

Qy   344 GGACAAGTACCGGCGCGCACCTGTCAGTGCTGCACCTGTGACCTCCTTTTGTCATCACACTCCCTT 403
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   342 GGACAAGTACAGGCGTGCACCTGTCAGTGCTGCACCTGTGACCTCCTTTTGTCATCACGCGTTCCCTT 401

Qy   404 CTGGGCCAGTTGATGCCATGGCCATGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   402 CTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCA 461

Qy   464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGCCTTCATCAGCCTGA 523
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   462 TGTCATCTACACAGTCAACCTCTACAGCAGTCCTCATCCTGCCTTCATCCATCAGTCTGA 521

Qy   524 CCGGTACCTCGCCATTGTCCACGCCACCACAGTCAAGGCCAAGGAAACTGCTGGCTGA 583
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   522 CCGGTACCTGGCCATCGTCCACGCCACCACAACAGTCAGAGGCCAAGGAGCGTGTTGGCTGA 581

Qy   584 AAAGGCAGTCTATGTGGGGCTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCAT 643
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   582 AAAGGTGTCTATGTGGCCGTCTGATCCCGCCCTCCTGCTGACTATTCCCGACTTCAT 641

Qy   644 CTTTGCCGACGTGCAGCTCAGCCAGGGGACATCAG 673
            ||||||  |||   || |||||||||||||||||
Db   642 CTTTGCCAACGTCAGTGAGGCAGATGACAG 671
```

RESULT 12
US-10-641-643-1213
; Sequence 1213, Application US/10641643

Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
;   Susan G. Stuart
;   Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
;   GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/641,643
;   FILING DATE: 14-Aug-2003
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: <Unknown>
;   FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1213:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1664 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g219868
; SEQUENCE DESCRIPTION: SEQ ID NO: 1213 :
US-10-641-643-1213

Query Match 65.1%; Score 445.8; DB 17; Length 1664;
Best Local Similarity 83.0%; Pred. No. 5.2e-132;
Matches 522; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

```
Qy    45 CAGGTGCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGT 104
            || |||||||||||||  |||||||  |||  || |||||||||||  ||||||||
Db     1 CGGCAGCAGGTAGCAAAGTGACAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCCACCATCTGG 60

Qy   105 AGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 164
            |||||||||||||||||||||||  |||||||||||||||||||||  ||||||||
Db    61 AGAACCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACC 115

Qy   165 GAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAAC 224
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   116 GAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGAACCCTGTTTCCGTGAAGAAAAT 175

Qy   225 GTCCATTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCATA 284
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   176 GCTAATTTCAATAAAATCTTCCTGCCACCATCTACTTCCATCATCTTCTTAACTGGCATT 235

Qy   285 GTCGGCAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGTAAGGAGCATGACG 344
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   236 GTGGGCAATGGGTTCATGGGTTCATGGGTTACCAGAAGAAGAAACTGAGAAGCATGACG 295

Qy   345 GACAAGTACCGGCGCGCACCTGTCAGTGCTGCACCTGTGACCTCCTTTTGTCATCACACTCCCTTC 404
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   296 GACAAGTACAGGCGTGCACCTGTCAGTGCTGCCGACCTCCTTTTGTCATCACGCGTTCCCTC 355
```

US-09-104-063-3

Query Match 64.1%; Score 439.2; DB 9; Length 1737;
Best Local Similarity 82.7%; Pred. No. 7e-130;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

```
Qy   50 GCAGTAGCAGTGACCCTGGAGGCGTTGGTGCTGCCGGTAACCACCACGGCTGTAGAGC 109
Db   20 GCGCGGCGCAAAGTGACGCGCGAGGGCTGGAGTGCTCCAGTGCTCCACCGCATCTGGAGAAC 79

Qy  110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTAATATACACTTCTGATAACTACACTCTGAAGA 169
Db   80 CAGCGGTTACCATGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134

Qy  170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGACCCTGCTTCCGGGGATGAAAACGTCCA 229
Db  135 AATGGGCTCAGGGACTATGACTCCATGAAGGACCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy  230 TTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTGACTCGGCATAGTCGG 289
Db  195 TTTCAATAAAATCTTCCTGCCACCATCTCCATCATCATCTTCTTAACTGGCATTGTGGG 254

Qy  290 CAATGGATTGGTGATCCTGTCATGGGTTACCAGAAGGAGTCTAAGGAGCATGACGGACAA 349
Db  255 CAATGGATTGGTCATCCTGGTCATCGGTTACCAGAGAGAAACTGAGAAGCATGACGGACAA 314

Qy  350 GTACCGGCTGCACCTGCCAGTGGCTGACCCTCCTGTTGTCATCACACTCCCTTCTCGGGC 409
Db  315 GTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTGTCATCACGCTCTCCCTTCTCGGGC 374

Qy  410 AGTTGATGGCCATGGCCTGACTCTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
Db  375 AGTTGATGCCGTGGCCAAACTCCTATGCAAGGCAGTCCATGTCAT 434

Qy  470 CTACACTGTCAACCTCTACAGCAGCGTTCATCCTGGCCTTCATCAGCCTGGACCCGGTA 529
Db  435 CTACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGCTA 494

Qy  530 CCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGC 589
Db  495 CCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGT 554

Qy  590 AGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTCCTGCTATACCTGACTTCATCTTTGC 649
Db  555 GGTCTATGTTGGCGTCTGGATCCCTGCCCCTCCTGCTATTCCCGACTTCATCTTTGC 614

Qy  650 CGACGTCAGCCAGGGGGACATCAG 673
Db  615 CAACGTCAGTGAGGCAGATGACAG 638
```

```
Qy  405 TGGGCAGTTGATGCCATGGCTGACTGTACTTGGGAAATTTTGTGTAAGGCTGTCCAT 464
Db  356 TGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGAACTTCCTATGCAAGGCAGTCCAT 415

Qy  465 ATCATTCACACTGTCAACCTCTACAGCAGCGTTCATCCTCGGCCCTTCATCAGCCTGGAC 524
Db  416 GTCATCACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCCTTCATCAGTCTGGAC 475

Qy  525 CGGTACCTCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAA 584
Db  476 CGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAA 535

Qy  585 AAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTCCTGCTATACCTGACTTCATC 644
Db  536 AAGGTGGTCTATGTTGGCGTCTGGATCCCTGCTGCCCCTCCTGCTGCTATTCCCGACTTCATC 595

Qy  645 TTTGCCGACGTCAGCCAGGGGGACATCAG 673
Db  596 TTTGCCAACGTCAGTGAGGCAGATGACAG 624
```

RESULT 13
US-09-104-063-3
; Sequence 3, Application US/09104063
Patent No. US2002016835A1
GENERAL INFORMATION:
APPLICANT: Lee, James
APPLICANT: Wood, William I.
TITLE OF INVENTION: PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/104,063
FILING DATE: 24-June-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/701265
FILING DATE: 22-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/664228
FILING DATE: 06-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear

RESULT 14
US-10-666-689-3
; Sequence 3, Application US/10666689
; Publication No. US20040037830A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Human PF4A Receptors, Nucleic Acid Encoding and
; TITLE OF INVENTION: Antibodies Binding Thereto
; FILE REFERENCE: P0706P2C2D2C1
; CURRENT APPLICATION NUMBER: US/10/666,689
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 09/104,063
; PRIOR FILING DATE: 1988-06-24
; PRIOR APPLICATION NUMBER: US 08/701,265
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/664,228
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/076,093
; PRIOR FILING DATE: 1993-06-11
; PRIOR APPLICATION NUMBER: US 07/810,782
; PRIOR FILING DATE: 1991-12-19

; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2143
; LENGTH: 1670
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L06797
US-09-880-107-2143

Query Match   63.9%;   Score 437.6;   DB 9;   Length 1670;
Best Local Similarity   83.7%;   Pred. No. 2.2e-129;
Matches 509;   Conservative   0;   Mismatches   94;   Indels   5;   Gaps   1;

Qy   66 CCTGAGGCGTTGGTGCTCCGTAACCACCACGGCTGTAGAGCCGAGTGTTGCCATGGAA 125
            ||||| ||||||||||| ||||| ||||||| || ||||| ||||||||
Db   18 CGCCGGCGGCCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAACCAGCGGTTACCATGA 77

Qy   126 CCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAGAGAAGTGGGGTCTGGAGAC 185
            ||||| |||| ||||||||||||||| ||||||||| ||| ||| || |||| ||||
Db   78 -----GGGGATCAGTATATACACTTCAGATAACTACACACCGAGGAAATGGGCTCAGGGGAC 132

Qy   186 TATGACTCCAACAAGGAGACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTC 245
            ||||| ||||| ||||||| ||| |||||| ||||| |||||| |||| ||||
Db   133 TATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTC 192

Qy   246 CTGCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATC 305
            ||||||||||| ||||||| ||| || |||||| |||||| |||||||| ||||||||
Db   193 CTGCCCACCATCTCATCCATCATCTTCTTAACTGGCATTGTGTGGGCAATGGATTGGTGATC 252

Qy   306 CTGGTCATGGGTACCAGAGGAGCTAAGGAGCCATGACGGACAAGTACCGGCTGCACCTG 365
            |||||||||||||||| |||||||| |||| ||||| ||||||| ||||||||||
Db   253 CTGGTCATGGGTACCAGCAGAGGAAACTGAGAAGCATGACGGACAAGTACCAGCTGCACCTG 312

Qy   366 TCAGTGGCTGACCTCCTCTTGTCATCACACTCCCCTCTCGGGCAGTGATGCCATGGCT 425
            ||||| ||||||||||| |||||||| ||||||||||||||| |||| ||||||
Db   313 TCAGTGGCCGACCTCCTCTTTGTCATCACGCTCCCCTTCTCGGGCAGTGATGCCGTGCCA 372

Qy   426 GACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTCACACTGTCAACCTC 485
            ||||||||||||||| |||||| | |||||| || |||| |||||||||| |||||||
Db   373 AACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTCACACAGTCAACCTC 432

Qy   486 TACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGCTACCTTCGCCCATTGTCCAC 545
            |||||| ||| ||||||||||||||||||||||||||| ||||| |||| ||||||||
Db   433 TACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTCAGTGTCGACCTCGGCCATCGTCCAC 492

Qy   546 GCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGAAAGGCAGTCTATGTGGGCGTC 605
            ||||||||||||| |||| ||||||||| || |||||||||||||| ||| |||||||
Db   493 GCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGTGGTCTATGTTGGCGTC 552

Qy   606 TGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACTTCATCTTTGCCAACTTC 665
            |||||||||||||||||| |||| ||||||||||||| ||||| |||||||||||||||| |
Db   553 TGGATCCCGCCCCTCCTGCTGCTGACTATTCCCGACTTCATCTTTGCCAACTTCATCTTTGC 612

Qy   666 GACATCAG 673
            || | ||
Db   613 GATGACAG 620

Search completed: May 17, 2004, 22:14:06
Job time : 337.298 secs

<hr/>

; PRIOR APPLICATION NUMBER: US 07/677,211
; PRIOR FILING DATE: 1991-03-29
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-666-689-3

Query Match   64.1%;   Score 439.2;   DB 13;   Length 1737;
Best Local Similarity   82.7%;   Pred. No. 7e-130;
Matches 516;   Conservative   0;   Mismatches   103;   Indels   5;   Gaps   1;

Qy   50 GCAGTAGCAGTGACCCCTCGAGGCGTTGGTGCTCCGTAACCACCACGGCTGTAGAGC 109
            |||| | ||| | |||| |||||||||||| ||||| |||||||||||
Db   20 GCGCGGCGCAAAGTGACGCCGAGGCCTGAGTGCTCCAGTAGCCACCACCGCATCTGGAGAAC 79

Qy   110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
            || ||||||| | ||||| |||||||||||||||||||| ||||||||| |||||
Db   80 CAGCGGTTACCATGA-----GGGGATCAGTATATATACACTTCAGATAACTACACACCGAGGA 134

Qy   170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAGACCCTGCTTCCGGGATGAAAACGTCCA 229
            |||| || || | ||||||||| ||||| ||||| | || |||||||| ||||| |
Db   135 AATGGGCTCAGGGACTATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAA 194

Qy   230 TTTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCGCATAGTCGG 289
            ||||||| |||||||||||| |||| ||| ||||||||||||| |||||| || | ||
Db   195 TTTCAATAAAATCTTCCTGCCCACCATCTCATCCATCATCTTCTTAACTGGCATTGTCATC 254

Qy   290 CAATGGATTGGTGATCCTGGTCATGGGTACCAGAGGAGCTAAGGAGCCATGACGGACGACAA 349
            |||||||||||||| ||||||||||||||||||||||||| ||||||||||||| |||||
Db   255 CAATGGATTGGTGATCCTGGTCATGGGTACCAGAGGAAACTGAGAAGCATGACGGACGACAA 314

Qy   350 GTACCGGCTGCACCTTGTCATCACCTTCTCGGGC 409
            |||| |||||||||| || ||||| |||||||
Db   315 GTACAGGCTGCACCTGTCAGTGGCCGACCTCCTCTTTGTCATCACGCTCCCCTTCTCGGGC 374

Qy   410 AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTCGTGTAAGGCTGTCCATATCAT 469
            |||||||||| | |||||||||||||||| |||| | | |||||| || ||||||
Db   375 AGTTGATGCCGTGCCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCAT 434

Qy   470 CTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGCTACCGGTA 529
            ||| | ||||||||||||||| ||| ||||||||||||||||||||| || | ||| |
Db   435 CTCACACAGTCAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTCAGTGTCGACCGCTA 494

Qy   530 CCTCGGCCATTGTGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGAAAGGC 589
            ||||||| |||| ||||||||||||| |||| |||||||| || ||||||||||
Db   495 CCTCGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGT 554

Qy   590 AGTCTATGTGGGCGTCTGGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGC 649
            ||||| ||| |||||||||||||||||||||||| |||| ||||||||||||| |
Db   555 GGTCTATGTTGGCGTCTGGATCCCGCCCCTCCTGCTGCTGACTATTCCCGACTTCATCTTTGC 614

Qy   650 CGACGTCAGCCAGGGGACATCAG 673
            || ||||||| || ||||||
Db   615 CAACGTCAGTGAGGCAGATGACAG 638

RESULT 15
US-09-880-107-2143
; Sequence 2143, Application US/09880107
; Patent No. US2002014981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14

OM nucleic - nucleic search, using sw model

Copyright GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 17, 2004, 05:50:31 ; Search time 1977.17 Seconds
(without alignments)
10345.919 Million cell updates/sec

Title: US-09-367-052-5
Perfect score: 685
Sequence: 1 ccatcctaatacgactcact.........gacatcagtcaggggggatga 685

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 1491090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
```
 1:  em_estba:*
 2:  em_esthum:*
 3:  em_estin:*
 4:  em_estmu:*
 5:  em_estov:*
 6:  em_estpl:*
 7:  em_estro:*
 8:  em_htc:*
 9:  gb_est1:*
10:  gb_est2:*
11:  gb_htc:*
12:  gb_est3:*
13:  gb_est4:*
14:  gb_est5:*
15:  em_estfun:*
16:  em_estom:*
17:  em_gss_hum:*
18:  em_gss_inv:*
19:  em_gss_prt:*
20:  em_gss_pln:*
21:  em_gss_fun:*
22:  em_gss_mam:*
23:  em_gss_mus:*
24:  em_gss_pro:*
25:  em_gss_rod:*
26:  em_gss_phg:*
27:  em_gss_vrl:*
28:  gb_gss1:*
29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	608.8	88.9	906	13	BU523025	BU523025 AGENCOURT
2	606.4	88.5	958	12	BG174412	BG174412 602334232
3	590	86.1	753	12	BG915636	BG915636 602814106
4	576.2	84.1	628	10	AW227957	AW227957 up19e03.y
5	544	79.4	955	12	BG173867	BG173867 602333928
6	536.4	78.3	581	14	CF540844	CF540844 UI-M-GW0-
7	511.4	74.7	585	13	BE627479	BE627479 uu52b07.y
8	503	73.4	504	13	BX513842	BX513842 BX513842
9	487.8	71.2	886	10	BF101953	BF101953 60175291 9
10	472.4	69.0	514	12	BG093377	BG093377 ut70g04.y
11	470.8	68.7	765	12	BM387369	BM387369 UI-R-CN1-
12	449.8	65.7	1201	13	BX418530	BX418530 BX418530
13	449.2	65.6	703	13	BI765768	BI765768 603046522
14	446.8	65.2	802	14	CB988971	CB988971 AGENCOURT
15	446.8	65.2	928	13	CD389217	CD389217 AGENCOURT
16	446.6	65.2	1201	13	BX358346	BX358346 BX358346
17	445.8	65.1	730	13	BX421959	BX421959 BX421959
18	445.8	65.1	1034	12	BM920800	BM920800 AGENCOURT
19	445.6	65.1	630	14	CB215245	CB215245 NISC_np03
20	445.6	65.1	702	12	BI761118	BI761118 603043593
21	445.6	65.1	910	12	BI762229	BI762229 603049139
22	445.6	65.1	913	12	BM051973	BM051973 603638877
23	445.6	65.1	950	13	BQ718617	BQ718617 AGENCOURT
24	445.2	65.0	658	12	BI835125	BI835125 603087792
25	445.2	65.0	742	12	BI597875	BI597875 603245357
26	444.8	64.9	728	9	AU117058	AU117058 AU117058
27	444.6	64.9	952	13	BX362477	BX362477 BX362477
28	444.4	64.9	1201	13	BX353253	BX353253 BX353253
29	444.2	64.8	945	13	BX443258	BX443258 BX443258
30	444.2	64.8	976	13	BU146456	BU146456 AGENCOURT
31	444	64.8	746	12	BI917014	BI917014 603177688
32	444	64.7	1201	13	BX462182	BX462182 BX462182
33	443	64.7	876	12	CA454255	CA454255 AGENCOURT
34	442.4	64.6	1201	13	BX421287	BX421287 BX421287
35	442.2	64.6	1201	13	BX462627	BX462627 BX462627
36	442	64.5	782	12	BG685901	BG685901 602638730
37	441.6	64.5	837	14	CB987769	CB987769 AGENCOURT
38	440.6	64.3	905	12	BI754094	BI754094 603027651
39	440	64.2	1051	12	BM545259	BM545259 AGENCOURT
40	439.4	64.1	1201	13	BX421511	BX421511 BX421511
41	438.6	64.0	911	13	BX443334	BX443334 BX443334
42	437.6	63.9	875	12	BI756157	BI756157 603030061
43	437.2	63.8	849	13	BU945383	BU945383 AGENCOURT
44	436.6	63.7	1201	13	BX396365	BX396365 BX396365
45	436.4	63.7	1091	13	BX397685	BX397685 BX397685

ALIGNMENTS

```
RESULT 1
BU523025
LOCUS       BU523025   906 bp   mRNA   linear   EST   13-SEP-2002
DEFINITION  AGENCOURT_10154335 NCI_CGAP_Co24 Mus musculus cDNA clone
            IMAGE:6529663 5', mRNA sequence.
ACCESSION   BU523025
VERSION     BU523025.1  GI:22833475
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14130 row: c column: 07
            High quality sequence stop: 668.
```

ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 958)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Gilbert Smith, Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM10254 row: o column: 07
           High quality sequence stop: 628.
FEATURES   Location/Qualifiers
    source   1..958
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="FVB/N"
           /db_xref="taxon:10090"
           /clone="IMAGE:4457694"
           /tissue_type="tumor, biopsy sample"
           /dev_stage="10 months, virgin"
           /lab_host="DH10B"
           /clone_lib="NCI_CGAP_Mam1"
           /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
           Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
           Library constructed by Life Technologies. Investigator
           providing samples: Gilbert Smith, NIH"
ORIGIN

Query Match   88.5%; Score 606.4; DB 12; Length 958;
Best Local Similarity 99.5%; Pred. No. 3.4e-136;
Matches 629; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
Qy    54 GTAGCAGTGACCCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGCGAGT 113
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GTAGCAGTGACCCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGCGAGT  60

Qy   114 GTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTG 173
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTG 120

Qy   174 GGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGTGGCATGGCATGGCCATTC 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGTGGCATGGCATGGCCATTC 180

Qy   234 AATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTGACTGGCATGGCATGGCGCAAT 293
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTGACTGGCATGGCATGGCGCAAT 240

Qy   294 GGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCATGACGGACAAGTAC 353
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCATGACGGACAAGTAC 300

Qy   354 CGGCTGCACCTGCAGTGGCTGACCTTTGGGAAATTTTGTTTGTAAGGCTGTCCATATCATCTAC 413
         ||||||||||||||||||||||||||||||||||||||||     ||||||||||||||||||
Db   301 CGGCTGCACCTGCAGTGGCTGACCTTTGGGAAA-TTTTGTTTGTAAGGCTGTCCATATCATCTAC 360

Qy   414 GATGCCATGGCAGCTGACTTTGGGAAATTTTGTGTCCGGCCTTCATCCTGGACCGGTACCTC 473
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GATGCCATGGCAGCTGACTTTGGGAAA-TTTTGTGTCCGGCCTTCATCCTGGACCGGTACCTC 419

Qy   474 ACTGTCAACCTCACAGCAGCGTTCTCATCCTGGCCTTCATCCTGGACCTGGACCTACCTC 533
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   420 ACTGTCAACCTCACAGCAGCGTTCTCATCCTGGCCTTCATCCTGGACCTGGACCTACCTC 479

Qy   534 GCCATTGTCACGCACCACAAACAGTCAACAGTCAAAGGCCAAGAAACTGCTGGCTGAAAAGGCAGTC 593
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

FEATURES   Location/Qualifiers
    source   1..906
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="FVB/N"
           /db_xref="taxon:10090"
           /clone="IMAGE:6529663"
           /lab_host="DH10B (T1 phage-resistant)"
           /clone_lib="NCI_CGAP_Co24"
           /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
           Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
           Average insert size 1.6 kb. Constructed by Life
           Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Query Match   88.9%; Score 608.8; DB 13; Length 906;
Best Local Similarity 98.7%; Pred. No. 8.6e-137;
Matches 626; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

```
Qy    52 AGTAGCAGTGACCCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGCGA 111
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGTACCAGTGACCCTGAGGCGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGCGA  60

Qy   112 GTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAG 171
         ||||||||||||||||||||||||||||||               ||||||||||||||||
Db    61 GTGTTGCCATGGAACCGATC-----AGTATATACACTTCTGATAACTACTCTGAAGAAG 114

Qy   172 TGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGTGGGATGAAAACGTCCATT 231
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   115 TGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGTGGGATGAAAACGTCCATT 174

Qy   232 TCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCTTTGACTGGCATGGCGGCA 291
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   175 TCAATAGGATCTTCCTGCCCACCATCATCTTCATCATCTTTGACTGGCATGGCGGCA 234

Qy   292 ATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGACATGACGGACAAGT 351
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   235 ATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAAGCTAAGGAGACATGACGGACAAGT 294

Qy   352 ACCGGCTGCACCTGTCAGTGGCTGACCTCCCTTTGTCATCACACTTCCCTTCTGGGCAG 411
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   295 ACCGGCTGCACCTGTCAGTGGCTGACCTCCCTTTGTCATCACACTTCCCTTCTGGGCAG 354

Qy   412 TTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCT 471
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   355 TTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCT 414

Qy   472 ACACTGTCAACCTCACAGCAGCGTTCTCATCCTGCCTTCATCAGCCTGACCGGTACC 531
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415 ACACTGTCAACCTCACAGCAGCGTTCTCATCCTGCCTTCATCAGCCTGACCGGTACC 474

Qy   532 TCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGAAACTGCTGGCTGAAAAGGCAG 591
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 TCGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGAAACTGCTGGCTGAAAAGGCAG 534

Qy   592 TCTATGTGGGCGTCTGGATCCGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCG 651
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535 TCTATGTGGGCGTCTGGATCCGATCCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCG 594

Qy   652 ACGTCAGCCAGGGGACATCAGTCAGGGGATGA 685
         ||||||||||||||||||||||||||||||||
Db   595 ACGTCAGCCAGGGGACATCAGTCAGGGGATGA 628
```

RESULT 2
BG174412
LOCUS   BG174412   958 bp   mRNA   linear   EST 06-FEB-2001
DEFINITION 602334232F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457694 5',
           mRNA sequence.
ACCESSION   BG174412
VERSION   BG174412.1 GI:12681115
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)

Db   480 GCCATTGTCCACGCCACCACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTC 539

Qy   594 TATGTGGGCGTCTGGATCCAGCCCTCCTGACTATACCTTGCCGAC 653
          |||||||||||||||||||||||||||||||||| ||||||||
Db   540 TATGTGGGCGTCTGGATCCAGCCCTCCTGA-TATACCTTCATCTTTGCCGAC 598

Qy   654 GTCAGCCAGGGGGACATCAGTCAGGGGATGA 685
          |||||||||||||||||||||||||||||||
Db   599 GTCAGCCAGGGGGACATCAGTCAGGGGATGA 630

```
RESULT 3
BG915636
LOCUS       BG915636              753 bp    mRNA    linear   EST 05-JUN-2001
DEFINITION  602814106F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4936465 5',
            mRNA sequence.
ACCESSION   BG915636
VERSION     BG915636.1  GI:14296112
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 753)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10868 row: 1 column: 02
            High quality sequence stop: 683.
FEATURES             Location/Qualifiers
     source          1..753
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="NMRI"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4936465"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam4"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Priscilla Furth,
                     NIH Reference for transgenic model: Li et al., Cell Growth
                     and Differentiation 7, 3-11 (1996)."
ORIGIN

Query Match      86.1%;   Score 590;  DB 12;  Length 753;
Best Local Similarity  98.7%;   Pred. No. 2.9e-132;
Matches  626;   Conservative  0;  Mismatches  5;   Indels  3;  Gaps  3;
```

Qy   50 GCAGGTAGCAGTGACCCTGAGGCGTTGGTGCTCCGGTACCACCACCACGGCTGTAGAGC 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10 GGAGGTAGCAGTGACCCTGAGGCGTTGGTGCTCCGGTACCACCACCACGGCTGTAGAGC 69

Qy   110 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70 GAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 129

Qy   170 AGTGGGGTCTGGAGACTATGAGCTCCAACAAGGAACCCTGCTTCCGGATGAAAAACGTCCA 229
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   130 AGTGGGGTCTGGAGACTATGAGCTCCAACAAGGAACCCTGCTTCCGGATGAAAAACGTCCA 189

Db   230 TTTCAATAGGATCTTCCTGCCGCCACCATCTACTTCATCATCTTCTTGACTGCATAGTCGG 289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   190 TTTCAATAGGATCTTCCTGCCGCCACCATCTACTTCATCATCTTCTTGACTGCATAGTCGG 249

Qy   290 CAATGGATTGGTGATCTGGTCATGGGTTACCAGAGAGAAGCTAAGGAGCATGACGGACAA 349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   250 CAATGGATTGGTGATCTGGTCATGGGTTACCAGAGAGAAGCTAAGGAGCATGACGGACAA 309

Qy   350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGC 409
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   310 GTACCGGCTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGC 369

Qy   410 AGTTGATGCCATGGCCTGACTGTACTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   370 AGTTGATGCCATGGCCTGACTGTACTTGGGAAA-TTCGTGTAAGGCTGTCCATATCAT 428

Qy   470 CTACACTGTCAACCTCTACAG-CAGCGTTCATCCTGGCCTTCATCAGCTGGACCGGT 528
          |||||||||||||||||||| ||| |||||||||||||||||||||| ||||||||
Db   429 CTACACTGTCAACCTCTACAGTCTGCCTTCATCCTGGCCTTCATCAGCCTGGACCGGT 488

Qy   529 ACCTCGCCATTGTCCACGCCACCACAGTCAAAGGCCAAGGAAACTGCTGGCTGAA-AAG 587
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   489 ACCTCGCCATTGTCCACGCCACCACAGTCAAAGGCCAAGGAAACTGCTGGCTGAACAAG 548

Qy   588 GCAGTCTATGTGGGCGTCTGGATCCAGCCCTCCTGACTATACCTGACTTCATCTTT 647
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   549 GCAGTCTATGTGGGCGTCTGGATCCAGCCCTCCTGACTATACCTGACTTCATCTTT 608

Qy   648 GCCGACGTCAGCCAGGGGGACATCAGTCAGGGGG 681
          ||||||||||||||||||||||||||||||||||
Db   609 GCCGACGTCAGCCAGGGGGACATCAGTCAGGGGGACATCAGTCAGGCGG 642

```
RESULT 4
AW227957
LOCUS       AW227957              628 bp    mRNA    linear   EST 10-DEC-1999
DEFINITION  up19e03.y1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:2654812 5'
            similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
            HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION   AW227957
VERSION     AW227957.1  GI:6557253
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 628)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            MGI:1032960
            Seq primer: -40RP from Gibco
            High quality sequence stop: 434.
FEATURES             Location/Qualifiers
     source          1..628
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="NMRI"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2654812"
```

REFERENCE   1 (bases 1 to 955)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/,
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Gilbert Smith, Ph.D.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10253 row: g column: 08
            High quality sequence stop: 652.

FEATURES   Location/Qualifiers
     source   1..955
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4457119"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP Mam1"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match   79.4%; Score 544; DB 12; Length 955;
Best Local Similarity   95.6%; Pred. No. 4.6e-121;
Matches 593; Conservative 0; Mismatches 20; Indels 7; Gaps 3;

Qy   73 GCGGTTGGTGCTCCGGTAACCACCACGGCTGTAGAGCGAGTGTTGCCATGGAACCGATCA 132
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1 GGCGTTTGGTGCTCCGGTAACCACCACGGCTGTGAGAGCGAGTGTTGCCATGGAACCGATCA 60

Qy   133 GTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCGGAGACTATGACT 192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCGGAGACTATGACT 120

Qy   193 CCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCA 252
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCA 180

Qy   253 CCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCA 312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCATCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCA 240

Qy   313 TGGGTTACCAGAAGAAGCTAAGGAGCTAAGGAGCATGACGGACAAGTACCACGGCTCAGTGG 372
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGGGTTACCAGAAGAAGCTAAGGAGCTAAGGAGCATGACGGACAAGTACCACGGCTCAGTGG 300

Qy   373 CTGACCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGT 432
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTGACCTCCTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGT 360

Qy   433 -ACTTTGGGAAATTTTTGTGTAAGGGTGTTAGGGTGGTCATTTTATTTATACATGTGAAACTTCAT 486
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AACTTTGGGAAATTTTTTGGGTTAGGGTGGTCATTTTATTTATACATGTGAAACTTCAT 420

Qy   487 ACAGCAGCGTTCTCATCCTGCCTTCATCAGCCTGACCGGTACCTCGGCCATTGTCCACG 546
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGAGAAGTGTTCATCCTGCCTTCATCAGCCTGACCGGTACCTCGGCCATTGTCCACG 480

Qy   547 CCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCATGTGGGGGTCT 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCATGTGGGGGTCT 540

Qy   607 GGATCCCAGCCCCTCCTGACTATACCTGACTTCATCTTTGCCGA-CGTCAGCCAGGGG 665
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||

Db   61 TGACTCCAACAAGGAACCCTGTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCT 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 248 GCCCACCATCACTTCATCATCTTCTTGACTCGGCATAGTCGGCAATGGATTGGTGATCCT 307
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GCCCACCATCTACTTCATCATCTTCTTCATCGTGCATAGTCGGCAATGGATTGGTGATCCT 180

Qy 308 GGTCATGGGTTACCAAGAAGCTAAGGAGCAGTGAGGAGGACAGTACCGGCTGCACCTGTC 367
        ||||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||
Db 181 GGTCATGGGTTACCAAGAAGCTAAGGAGCATGACGGACGAACAAGTACCGGCTGCACCTGTC 240

Qy 368 AGTGGCTGACCTCCTCTTGTCATCACACTCCCTTCTGGGCAGTGATGCCATGGCTGA 427
        |||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db 241 AGTGGCTGACCTCCTCTTGTCATCACACTCCCTTCGGGCAGTTGATGCCATGGCTGA 300

Qy 428 CTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCCTA 487
        |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 301 CTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCCTA 360

Qy 488 CAGCAGGGTTCTCATCCTCGGCCTTCATCAGCCTGGACCGTACCTCCGCCATTGTCCACGC 547
        || ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db 361 CAGCAGGGTTCTCATCCTCGGCCTTCATCAGCCTGGACCGTACCTCCGCTATTGTCCACGC 420

Qy 548 CACCAACAGTCAAAGGCCAAGAAACTGCTGCGTGCTGANAAGGCAGTCATGTGGGCGTCTG 607
        |||||| |||||| ||||||||||||||||||||| ||||| ||||||||||||||||||
Db 421 CACCAACAGTCAGAGGCCAAGGAAACTGCTGGCTGANAAGGCAGTCTATGTGGGCGTCTG 480

Qy 608 GATCCCAGCCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGA 667
        || ||||||||||| |||||||||||||||||||||||| |||||||||||||||||
Db 481 GAT-CCAGCCCCTCCTGACTATACCTGACTTCATCTTTNGCGACGTCAGCCAGGGGGA 539

Qy 668 CATCAGTCAGGGGATGA 685
        ||||||||||||||||
Db 540 CATCAGTCANGGGATGA 557

RESULT 7
BE627479
LOCUS   BE627479   585 bp   mRNA   linear   EST 24-AUG-2000
DEFINITION   uu52b07.y1 Soares_thymus_2NbMT Mus musculus cDNA clone
             IMAGE:3375541 5' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
             RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION   BE627479
VERSION   BE627479.1 GI:9907901
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 585)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:1085145
             Seq primer: -40RP from Gibco
             High quality sequence stop: 463.
FEATURES   Location/Qualifiers
     source   1..585
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3375541"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_thymus_2NbMT"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified

Db 541 GGATCCCAGCCCCTCCTGACTATACCTGACTTCATCTTGCCGACCGTCAGCCAGGGG 600
        |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Qy 666 GACATCAGTCAGGGGATGA 685
        ||||||||||||||||
Db 601 GACATCAGTCAGGGGATGA 620

RESULT 6
CF540844
LOCUS   CF540844   581 bp   mRNA   linear   EST 12-SEP-2003
DEFINITION   UI-M-GW0-cio-1-17-0-UI.r1 NIH_BMAP_GW0 Mus musculus cDNA clone
             IMAGE:30542392 5', mRNA sequence.
ACCESSION   CF540844
VERSION   CF540844.1 GI:34593367
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 581)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
             http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
             (BMAP)
             Seq primer: pYX-5.
FEATURES   Location/Qualifiers
     source   1..581
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30542392"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 15.5,16.5,17.5,18.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GW0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site 1: EcoR I;
                     Site 2: Not I; The library was constructed according
                     BonaTdo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NoI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CTGCGCCTC. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP); 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."

ORIGIN
Query Match   78.3%; Score 536.6; DB 14; Length 581;
Best Local Similarity   98.6%; Pred. No. 2.3e-119;
Matches 550; Conservative   0; Mismatches   7; Indels   1; Gaps   1;

Qy 128 GATCAGTGTGAGTAGTATATACACTTCTGATAACTAACTACTCTGAAGAAGTGGGGCTGGAGAGACTA 187
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GATCAGTGTGAGTAGTATATACACTTCTGATAACTAACTACTCTGAAGAAGTGGGTTCTGGAGAGACTA 60

Qy 188 TGACTCCAACAAGGAACCCTGTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCT 247
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

FEATURES   Location/Qualifiers
  source   1..504
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="NMRI"
                  /db_xref="taxon:10090"
                  /clone="IMAGp9980056625 ; IMAGE:2654812"
                  /tissue_type="tumor, gross tissue"
                  /dev_stage="5 months"
                  /lab_host="DH10B"
                  /clone_lib="NCI_CGAP_Mam4"
                  /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                  Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT
                  Library constructed by Life Technologies. Investigators
                  providing samples: Lothar Hennighausen/Priscilla Furth,
                  NIH Reference for transgenic model: Li et al., Cell Growth
                  and Differentiation 7, 3-11 (1996)."

ORIGIN

Query Match   73.4%;   Score 503; DB 13;   Length 504;
Best Local Similarity   99.8%;   Pred. No. 3e-111;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps   0;

Qy   50 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACCGGCTGTAGAGC 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCAGGTAGCAGTGACCCTCTGAGGCGTTGGTGCTCCGGTAACCACCACCGGCTGTAGAGC   60

Qy   110 GAGTGTTGCCATGGAACGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GAGTGTTGCCATGGAACGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 120

Qy   170 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 229
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGTGGGGTCTGGAGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCA 180

Qy   230 TTTCAATAGGATCTTCCTGCCCACCATCACTTCTTCTCATCATCTTCTTGACTGGCATAGTCGG 289
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TTTCAATAGGATCTTCCTGCCCACCATCACTTCTTCTCATCATCTTCTTGACTGGCATAGTCGG 240

Qy   290 CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAGAGTCAAGGAGNTAAGGAGNTAAGGAGCATGACGACGACGACAA 349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAATGGATTGGTGATCCTGGTCATGGGTTACCAGAAGAGAGTCAAGGAGNTAAGGAGCATGACGACGACGACGACAA 300

Qy   350 GTACCGGCTGCACCTGTGACTGGCTGACCTCCTTTGTCATCACACTCCCTTCTGGGC 409
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GTACCGGCTGCACCTGTGACTGGCTGACCTCCTTTGTCATCACACTCCCTTCTGGGC 360

Qy   410 AGTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCAT 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCAT 420

Qy   470 CTACACTGTCAACCTCTACAGCAGCGTTCATCCTGCCTTCATCAGCCTGGACCGGTA 529
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTACACTGTCAACCTCTACAGCAGCGTTCATCCTGCCTTCATCAGCCTGGACCGGTA 480

Qy   530 CCTCGCCATTGTCCACGCCACCAA 553
          ||||||||||||||||||||||||
Db   481 CCTCGCCATTGTCCACGCCACCAA 504

RESULT 9

ORIGIN

Query Match   74.7%;   Score 511.4; DB 10;   Length 585;
Best Local Similarity   96.8%;   Pred. No. 3e-113;
Matches 521; Conservative 0; Mismatches 17; Indels 0; Gaps   0;

Qy   121 TGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTG 180
          |||||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db   48 TGTCCCTGTTTCCTTTGCAGATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTTCTG 107

Qy   181 GAGACTATGACTCCAACAAGGAACCCGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 240
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   108 GAGACTATGACTCCAACAAGGAACCCGCTTCCGGGATGAAAACGTCCATTTCAATAGGA 167

Qy   241 TCTTCCTGCCCACCATCACTTCTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 300
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   168 TCTTCCTGCCCCACCATCACTTCTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGG 227

Qy   301 TGATCCTGGTCATGGGTTACCAGAAGAGCATGACGACGACGACAAGTACCGGCTGC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   228 TGATCCTGGTCATGGGTTACCAGAAGAGCATGACGACGACGACAAGTACCGGCTGC 287

Qy   361 ACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCTTCTGGGCAGTTGATGCCA 420
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   288 ACCTGTCAGTGGCTGACCTCCTTTGTCATCACACTCCCTTCTGGGCAGTTGTTGCCA 347

Qy   421 TGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCATCACACTGTCA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   348 TGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCATCACACTGTCA 407

Qy   481 ACCTCTACAGACAGCGCTTCTCATCCTGGCCTTCAGCCTCGGACCGGTACCTCGGCCATTG 540
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   408 ACCTCTACAGCAGCGTTCTCATCCTGGCCTTCAGCCTCGGACCGGTACTTCGGCCTATTG 467

Qy   541 TCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTCATGTGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   468 TCCACGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTCATGTGG 527

Qy   601 GCGTCTGGATCCCAGCCCTCCTGACTATACCTTGCCGACGTCAG 658
          ||||||||||||||||||||||||||||||| |||||||||||
Db   528 GCGTCTGGATCCCAGCCCTCCTGACTATACCTTGACTTCATCNTGCCGACGTCAG 585

RESULT 8
BX513842

LOCUS   BX513842   504 bp   mRNA   linear   EST 25-JUN-2003
DEFINITION BX513842 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGp9980056625 ;
            IMAGE:2654812, mRNA sequence.
ACCESSION   BX513842
VERSION   BX513842.1 GI:32243853
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 504)
  AUTHORS   Heil,O., Eberl,L., Neubert,P., Peters,M., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE   Mouse UnigeneSet - R2PD2
  JOURNAL   Unpublished (2003)
  COMMENT   Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp9980056625.

BF101953
LOCUS   BF101953   886 bp   mRNA   linear   EST 19-OCT-2000
DEFINITION 601752919F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3980525 5',
            mRNA sequence.
ACCESSION   BF101953
VERSION   BF101953.1 GI:10884479
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 886)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9176 row: e   column: 06
            High quality sequence stop: 665.
FEATURES   Location/Qualifiers
     source   1..886
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3980525"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam1"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                     Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
ORIGIN

Query Match   71.2%;   Score 487.8; DB 10; Length 886;
Best Local Similarity 98.7%; Pred. No. 1.9e-107;
Matches 534; Conservative 0; Mismatches 2; Indels 5; Gaps 4;

Qy   77 TTGGTGTCCGTAACCACCACGGCTGTAGAGCGAGTGTTGCCATGGAACCGATCAGTGT 136
         ||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||
Db   7 TTGGTGTCCGGTAACCACCACGGCTGTGTAGAGCGAGTGTTGCCATGGAACCGATCAGTGT 66

Qy 137 GAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAA 196
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 GAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAA 126

Qy 197 CAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCAT 256
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 127 CAAGGAACCCTGCTTCCGGGATGAAAACGTCCA-TTCAATAGGATCTTCCTGCCCACCAT 185

Qy 257 CTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGG 316
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 186 CTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGG 245

Qy 317 TTACCAGAAGAAGTAAGGAGCTAAGGACATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGA 376
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 246 TTACCAGAAGAAGTAAGGAGCTAAGGACATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGA 305

Qy 377 CCTCCTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACTT 436
         ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 306 CCTCCTC-TTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACTT 364

Qy 437 TGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGT 496
         || || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 365 --GGGAAATTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGT 422

Qy 497 TCTCATCTGGCCTTCATCAGCCTGGACCGGACCGGTCCAGGCCACCAACAG 556
         ||||||||||||||||||||||||||||||| |||||||||||||||||||
Db 423 TCTCATCTGGCCTTCATCAGCCTGGACCGGTACCCGCCACCAACAG 482

Qy 557 TCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGGTCTGATCCCAGC 616
         |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db 483 TCAAAGGCCAAGGAAACTG-TGGCTGAAAAGGCAGTCTATGTGGGGTCTGATCCCAGC 541

Qy 617 C 617
         |
Db 542 C 542

RESULT 10
BG093377
LOCUS   BG093377   514 bp   mRNA   linear   EST 26-JAN-2001
DEFINITION ut70g04.y1 Soares mouse NMGB bcell Mus musculus cDNA clone
            IMAGE:3333318 5' similar to SW:CCR4_MOUSE P70658 C-X-C CHEMOKINE
            RECEPTOR TYPE 4 ;, mRNA sequence.
ACCESSION   BG093377
VERSION   BG093377.1 GI:12575940
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 514)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1077482
            Trace considered overall poor quality
            Seq primer: -40RP from Gibco
            High quality sequence stop: 1.
FEATURES   Location/Qualifiers
     source   1..514
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3333318"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Soares_mouse_NMGB_bcell"
                     /note="Organ: germinal B-cell; Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site 1: Not I;
                     Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCGGCCGCTGTTTTTTTTTTTTTTTTTTTTT
                     T 3'], double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized; constructed by Bento Soares and
                     M.Fatima Bonaldo."
ORIGIN

Query Match   69.0%;   Score 472.4; DB 12; Length 514;
Best Local Similarity 94.9%; Pred. No. 8e-104;
Matches 488; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 122 GGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGG 181
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGG 60

Qy 182 AGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGAT 241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGACTATGACTCCAACAAGGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGAT 120

```
Qy   242  CTTCCTGCCCACCATCACTTCATCATCTTCTGACTGGCATAGTCGGCAATGGATTGGT  301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CTTCCTGCCCACCATCACTTCATCATCTTCTGACTGGCATAGTCGGCAATGGATTGGT  180

Qy   302  GATCCTGGTCATGGGTTACCAGAAGAGAGCTAAGGAGCATGACGGACAAGTACCGGCTGCA  361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GATCCTGGTCATGGGTTACCATAATA:GCTAAGGAGCATGACGGACAAGTACCGGGCTGTA  240

Qy   362  CCTGTCAGTGGCTGACCTCCCTTGGTCATCACACACTCCCCTTCTGGGCAGTTGATGCCAT  421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  CCTGTCAGTGGCTGACCTCCCTTTGTCATCACACACTCCCCTTCTGGGCAGTTGATGCCAT  300

Qy   422  GGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCGTGTCCATATCATCTCACACTGTCAA  481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GGCTGACTGGTACTTTGGGAAATTTGTGTGTAAGGCGTGTCCATATCATCATCACACTGTCTA  360

Qy   482  CCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGCCTGGACCAGTTACCTTCGGCCATTGT  541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CCTCTACAGCAGCGTTCTCATCATCGTCCTTCATCAGCCTGTACCGGTACCTTCGACATTGT  420

Qy   542  CCACGCCACCACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGG  601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GCACGCCTACCACAGTCATAGGCCAATGTAACTGCTGGCTGAATACGCAGTCTATGTCGG  480

Qy   602  CGTCTGGATCCCAGCCCTCCCTCCTGACTATACCT  635
          ||||||||||||||||||||||||| |||||||||
Db   481  CGTCTGGATCCCCAGCCCATCCTCCTGACTATATCT  514
```

RESULT 11
BM387369

LOCUS   BM387369   765 bp   mRNA   linear   EST 17-JAN-2002
DEFINITION UI-R-CN1-cjj-g-19-0-UI.s1 UI-R-CN1 Rattus norvegicus cDNA clone
            UI-R-CN1-cjj-g-19-0-UI 3', mRNA sequence.
ACCESSION   BM387369
VERSION   BM387369.1 GI:18187422
KEYWORDS   EST.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 765)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE   Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED   8889548
COMMENT   Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to identify it as a clone from the
            normalized cervix library cDNA Library Preparation: M.B. Soares Lab
            Clone distribution: clones will be available through Research
            Genetics (www.resgen.com)
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1..765
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-CN1-cjj-g-19-0-UI"

                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="UI-R-CN1"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CN1
                     library is a subtracced library derived from the following
                     pool of seven normalized rat libraries: normalized rat
                     seminal vesicles, normalized rat penis, normalized rat
                     bladder, normalized rat cervix, normalized rat brown
                     adipose, normalized rat fundus, and normalized rat
                     salivary gland. It was constructed according to the
                     procedure described by Bonaldo, Lennon & Soares (Genome
                     Research Genome 6: 791-806, 1996). For construction of
                     the CN1 library, plasmid DNA from the pool of normalized
                     libraries was electroporated into competent bacteria for
                     the production of single-stranded circular DNA. This was
                     then used as a tracer in a subtractive hybridization with
                     a driver (PCR amplified inserts from a plasmid DNA template
                     preparation) comprising: a) a pool of about 34,000 clones
                     from the Rat Unigene Set corresponding to plates R-5-AA-NN
                     excluding plates R-5-MM and MN. This pool represented 40%
                     of the final driver population. b) a pool of about 29,000
                     clones from subtracted libraries CA0 and CA1 corresponding
                     to plates R-CA0-AWV through R-CA0-AXS, R-CA0-AZX through
                     R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS,
                     R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through
                     R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA,
                     R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through
                     R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA1-BBA through
                     R-CA1-BDA, R-CA1-BHZ through R-CA1-BJF, R-CA1-BJR,
                     R-CA1-BJT through R-CA1-BKB, R-CA1-BKD, R-CA1-BKF,
                     R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through
                     R-CA1-BLN, R-CA1-BLS, R-CA1-BLU-V, R-CA1-BNR, and
                     R-CA1-BLE. The resulting pool represented 20% of the
                     final driver population. c) a pool of about 15,000 clones
                     from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s
                     and normalized libaries CS0, CT0, CU0, CW0, and CX0
                     corresponding to plates R-CS0s-CBD through R-CS0s-CBO,
                     R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through
                     R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN
                     through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BUP,
                     R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP,
                     R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVV through
                     R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through
                     R-CX0-BXM. The resulting pool represented 5% of the final
                     driver population. d) a pool of about 5,000 clones (1,000
                     from non-normalized eye library CV0 and 4,000 from
                     normalized eye library CV1) corresponding to plates
                     R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC,
                     R-CV1-SSE through R-CV1-BTC, and R-CV1-BVO through
                     R-CV1-BVU. This pool represented about 5% of the final
                     driver population. e) A pool of about 10,000 clones from
                     subtracted library BS2, BV0 and BV0p (7-9.5 kb cDNA
                     library fraction from rat whole embryo), and BX0 (0.5-7kb
                     cDNA library fraction from rat whole embryo) corresponding
                     to plates R-BS2-BDB through R-BS2-BFP, R-BV0-ANK through
                     R-BV0-ANR, R-BV0p-AOI through R-BV0p-AOX, and R-BX0-AQY
                     through R-BX0-ASH. The resulting pool represented 5% of
                     the final driver population. f) a pool of about 7,000
                     clones from the seven non-normalized libraries that make
                     up the tracer including CV0, CZ0, DA0, DB0, DC0, DD0, and
                     DE0 corresponding to plates R-CV0-BXP through R-CV0-BX2,
                     R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ
                     through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ
                     through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DC0-CAY
                     through R-DC0-CBA, R-DD0-BZR through R-DD0-CAL.
                     R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The
                     resulting pool represented about 10% of the final driver
                     population. g) a pool of about 2,000 clones from the pool
                     of normalized libraries, CN0, that makes up the tracer.
                     The corresponding plates are R-CN0-BKW through R-CN0-BLD,
                     R-CN0-BLG, R-CN0-BLP through R-CN0-BLR, R-CN0-BLT,
                     R-CN0-BLW-X, R-CN0-BMB, and R-CN0-BMF through R-CN0-BML.

This pool represented 5% of the final driver population.
h) a pool of the 28 most abundant clones in the CNO pool
corresponding to the following addresses: bkw-a-09-0-UI,
bkw-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI,
bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI,
bkx-a-11-0-UI, bky-a-05-0-UI, bkz-a-06-0-UI,
bkz-a-11-0-UI, bkz-c-06-0-UI, bkz-c-09-0-UI,
bkz-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI,
bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI,
blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI,
blc-e-95-0-UI, bld-1-08-0-UI, bld-f-02-0-UI,
blq-a-04-0-UI, blr-a-05-0-UI, bit-f-08-0-UI. This pool
represented 5% of the final driver population. i) One
abundant CNO clone (corresponding to the address
bkz-a-11-0-UI) was digested with Not I and Eco RI and the
resulting insert was gel purified. This purified insert
was added directly to the driver so that it represented 5%
of the final driver population.
TAG TISSUE=cervix
TAG_LIB=UI-R-CN1
TAG_SEQ=GACCA"

ORIGIN

Query Match
Best Local Similarity 68.7%;   Score 470.8;   DB 12;   Length 765;
Matches 496; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy   136 TGAGTATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCA 195
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   105 TGGAAATATACACTTCGGATAACTACTCCGAAGAAGTAGGGTCTGGAGACTATGACTCCA 164

Qy   196 ACAAGGAACCCTGCTTCCCGGATGAAAAGTCCATTCAATAGATCTTCCTGCCCACCA 255
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   165 ACAAGGAACCCTGCTTCCCGGATGAAAAGTCCATTCAATAGATCTTCCTGCCCACCA 224

Qy   256 TCTACTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGAATGGTGATCGGTCATGG 315
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   225 TCTATTTTATTATCTTCTTGACTGGCATAGTGGCATAGTGGGTAATCCTGGTCATGG 284

Qy   316 GTTACCAGAAGAAGTAAGGACGCATGAGGACGAAGTACCGGCTGCACTGGCTG 375
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   285 GTTACCAGAAGAAGTGAGGAGAGCATGACAGAGAAGTACCGGCTGCACTGGCTG 344

Qy   376 ACCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACT 435
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   345 ACCTCCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACT 404

Qy   436 TTGGGAAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTTCAACCTCTACAGCAGCG 495
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   405 TTGGGAAATTTTTATGTAAGGCTGTGCATATCATCTACACCGTCAACCTTTACAGCAGTG 464

Qy   496 TTCTCATCCTCGGGCCTTCATCAGCCTGGACCTGGCTCCGCCATTGTCCACGCCACCACA 555
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   465 TTCTCATCCTCGGGCCTTCATCAGCCTGGACCTGCTACCTTGCCATTGTCCACGCCACCACA 524

Qy   556 GTCAAAGGCCAAGGAAACTCTGGCTGCTGAAAAGGCAGTCTATGTGGGGGCTCTGGATCCCAG 615
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   525 GCCAGAGCGCGAGGAAGCTGTCGTGCTGAAAAGGCGTCTTATGTGGTCTGGATCCCCG 584

Qy   616 CCCTCCCTCCTGACTATACCCGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAG 673
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   585 CCCTCCCTCCTGACTATCCCTGACATCATCTTCGCCGATGTCAGCCAGGGGGACGGCAG 642

ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 1272.r For
             more information about this cluster, see
             http://www.genoscope.cns.fr/
             cgi-bin/cluster.cgi?seq=CS0DF006DE08QP1&cluster=1272.r. Contact :
             Feng Liang Email : fliang@lifetech.com URL:
             http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0DF006DE08QP1.
FEATURES   Location/Qualifiers
     source   1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF006YJ16"
                     /tissue_type="FETAL BRAIN"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

Query Match   65.7%;   Score 449.8;   DB 13;   Length 1201;
Best Local Similarity 82.8%;   Pred. No. 3.6e-98;
Matches 526; Conservative 1; Mismatches 103; Indels 5; Gaps 1;

Qy   39 CCCGGCCAGGTCGCAGGTAGCAGTGAGTCCTCCGGTTGGTGCTAACCACCAC   98
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCCGGGATGCAGCAGGTAGCAAGTGACGCCCGAGGGCCGAGTGCTCCAGTGCCCACCGC 120

Qy   99 GGCTGTAGAGCGAGTGTTGCCATGGCAGTTGATGCTGAGTATATACACTTCTGATAAC 158
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ATCTGGAGAACCAGCGGTTACCATGGAGGGGA-----TCAGTATATACACTTCAGATAAC 175

Qy   159 TACTCTGAAGAGTGGGGTCTGGAGACTATGACTCCAACAACAAGGAACCCTGCTTCCGGAT 218
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   176 TACACCGAGAAATGGGCTCAGGGGACTATGAYTCCATGAAGGGGACACCGTGTTTCCGTGAA 235

Qy   219 GAAAAGTCCATTCAATAGGATCTTCCTGCCCACCATCTACTTCTTCTTCTTGACT 278
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   236 GAAATGCTAATTCAATAAAATCTTCCTGCCCACCATCTACTTCTTCTTCTTTTAACT 295

Qy   279 GGCATAGTCGGCAATGATGGTGATCCTGGTCATGGGTTACCAGAAGAAGTAAGGAGC 338
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   296 GGCATTGTGGGCAATGATTGGTCATCCTGGTTCATGGGTTACCAGAAGAAACTGAGAAGC 355

Qy   339 ATGACGGACAAGTACCGGCTGCACTGGCTGACACTGTACTTTGGGAAATTTTGTGTAAGGCT 398
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   356 ATGACGGACAAGTACAGGCTGCACCTGTCAGTGGTACTTCTTTGGGAACTTCCTATGCAGGCA 415

Qy   399 CCCTTCTGGGCAGTTGATGCCATGGCCATGGCTGACTGGTACTTTGGGAAATTTTTGTGTAAGGCT 458
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   416 CCCTTCTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGAACTTCCTATGCAGGCA 475

Qy   459 GTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCCTGGCCTTCATCAGC 518
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   476 GTCCATGTCATCTACACAGTCAACCTCTACACAGCAGTGTCCTCACCCTGGCCTTCATCAGT 535

Qy   519 CTGGACCGGTACCTCGCCATTGTCCACGCCACCACAGTCAAAGGCCAAGAGTCAAAGGCCCAAGGGAAACTGCTG 578
           |||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   124 TACACTTCAGATAACTACACCGAGAGAAATGGGCTCAGGGACTATGATGACTCCATGAAGGAA 183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   204 CCTGCTTCCGGATGAAAACGTCCATTCAATTCATAGAGATCTTCCTGCCACCATCTACTTC 263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 CCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTACTCC 243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   264 ATCATCTTCTTGACTGCCATAGTGCGGCAATGAGTTGGTGATCCTGGGTCATGGGTTACCAG 323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 ATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTTCATGGGTTACCAG 303
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   324 AAGAAGTAAGGAGCATGACGGACAAGTACCGGCTGCACCGTGCACCTGGCTGCTGACCTCCTC 383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   304 AAGAAAACTGAGAGCATGACGGACAAGTACACGGCTGCACCGTGCACCTGGCTGCCGACCTCCTC 363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   384 TTTGTCATCACACTCCCCTTCTGGCAGTTGATGCCATGGCTGACTGGTACTTTGGGAAA 443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 TTTGTCATCACGCTTCCCTTCTGGCAGTTGATGCCATGGCTGGCAAACTGTACTTTGGGAAC 423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   444 TTTTTGTTGTAAGGCTGTCCATATCACACTGTCAACCTCTACAGACGCGTTCTCATC 503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 TTCCTATGCAAGGCAGTCCATGTCATCACACAGTCAACCTCTACACAGCAGTGTCCTCATC 483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   504 CTGGCCTTCATCAGCCTGGACCGGTACCTGCCCATTGTCCACGCCACCAAACAGTCAAAGG 563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 CTGGCCTTCATCATCTGGACCGGTACCTGCCCATCGTCCACGCCACCAACAGTCAGAGG 543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   564 CCAAGGAAACTGCTGGCTGAAAAAGGCAGTCTATGTGGGGCGTCTGGATCCCGATCCCCAGCCCTCTC 623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 CCAAGGAAGCTGTTGGCTGAAAAAGGTGGTCTATGTTGGCGTCTGGATCCCGTCTGGATCCCCTCTG 603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   624 CTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAG 673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   604 CTGACTATTCCGGACTTCATCTTTGCCAACGTCAGTGAGGCCAGTGACACAG 653
```

RESULT 14
CB988971
LOCUS   CB988971   802 bp   mRNA   linear   EST 01-MAY-2003
DEFINITION AGENCOURT 13893109 NIH_MGC_147 Homo sapiens cDNA clone
            IMAGE:30346124 5', mRNA sequence.
ACCESSION   CB988971
VERSION   CB988971.1 GI:30283491
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 802)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM385 row: b column: 21
            High quality sequence stop: 663.
FEATURES   Location/Qualifiers
     source   1..802
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30346124"
                     /tissue_type="Human Placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_147"

```
Db   536 CTGGACCGCTACCTGCCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTG 595
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   579 GCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCCTCCTGACTATACCTGAC 638
             ||||||  |||||||  ||||| |||||||||||||||  |||||
Db   596 GCTGAAAAGGTGGTCTATGTTGGCGTCTGGATCCCGTCTGGATCCCTCCCGAC 655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   639 TTCATCTTTGCCGACGTCAGCCAGGGGGACATCAG 673
             ||||||  ||||||||||||||  |||||
Db   656 TTCATCTTTGCCAACGTCAGTGAGGCCAGTGACAGATGACAG 690
```

RESULT 13
BI765768
LOCUS   BI765768   703 bp   mRNA   linear   EST 25-SEP-2001
DEFINITION 603046522F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186837 5',
            mRNA sequence.
ACCESSION   BI765768
VERSION   BI765768.1 GI:15757346
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 703)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1466 row: 1 column: 06
            High quality sequence start: 24
            High quality sequence stop: 703.
FEATURES   Location/Qualifiers
     source   1..703
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5186837"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female. Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon cloning). Average insert size 1.4 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber (Invitrogen). Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
ORIGIN

Query Match   65.6%; Score 449.2; DB 12; Length 703;
Best Local Similarity 81.8%; Pred. No. 3.9e-98;
Matches 532; Conservative   0; Mismatches 113; Indels   5; Gaps   1;

Qy   24 GGGGTCGAGCGGCGCCCCCGGCAGGTAGCAGTGCAGGTAGCAGTGACCCTTCTGAGGCCGTTGGTGC 83
             |||||||||| ||||||||||||| ||||||| |||||||||||||||||||||||||
Db   9 GGTCCGAGTCGTTGGATCGTGCGGCCAGCAGGTAGCAAAGTGACCGGCGAGGGGCCTGAGTGC 68
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   84 TCCGGTAACCACCACCGCTGTGAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATA 143
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   69 TCCAGTAGCCACCGCCACCGCCATCTGGAGAACCAGCGGTTACCATGGAGGGGA----TCAGTATA 123
             |||||||||||||||||||||||||||||||||||||||| ||||||||
Qy   144 TACACTTCTGATAACTACTCTGAGAAGTGGGGTCTGAGACTATGACTCCAACAAGGAA 203
```

```
                    /note="Organ: placenta; Vector: pBluescriptR; Site 1:
                    all-XhoI; Site 2: BanH; Oligo-dT primed using primer
                    5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average
                    insert size 2.3 kb and normalized to ROT 5. This is a
                    primary library enriched for full-length clones and
                    constructed using the Cap-trapper method (Carninci, in
                    preparation). Library constructed by M. Brownstein
                    (NIMH/NHGRI, National Institutes of Health). Note: This is
                    a NIH_MGC library."

ORIGIN

       Query Match         65.2%;    Score 446.8; DB 14;  Length 802;
       Best Local Similarity   83.0%;    Pred. No. 1.6e-97;
       Matches 523; Conservative    0; Mismatches 102; Indels    5; Gaps    1;

Qy     44 GCAGGTGCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
          ||||||||||||||||||  | |||||| |||||||||||||||| ||||||||||||||
Db     40 GCGGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 99

Qy    104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
          ||||||||||||||||||||||||||||||||     |||||||||||||||||||||||
Db    100 GAGAACCAGCGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACAC 154

Qy    164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGACACCCGCTTCCGGGATGAAAA 223
          |||||||||||||||  ||||||||||||  |||||||||||||||||||||||||||||
Db    155 CGAGGAAATGGGGCTCAGGGGACTATGACTCCATGAAGGAACCCGTTTCCGTGAAGAAAA 214

Qy    224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCACTCATCATCATCTTAACTGGCAT 283
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db    215 TGCTAATTTCAATAAAAATTCCTGCCCACCATCTACTTCACTCATCATCATCTTAACTGGCAT 274

Qy    284 AGTCGGCAATGGATTGGTCATCCTGGTCATGGGGTTACCAGAAGAAGTAAGGAGCATGAC 343
          ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    275 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGGTTACCAGAAGAAAC TGAGAAGCATGAC 334

Qy    344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACTCCTCTTGTCATCACACTCCCCTT 403
          |||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db    335 GGACAAGTACAGGCTGCACCTGTCAGTCAGTGGCCGACTCCTCTTGTCATCACGGCTTCCCTT 394

Qy    404 CTGGGCAGTTGATGCCATGGCTGACTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
          ||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db    395 CTGGGCAGTTGATGCCGTGGCAAACTGTACTTTGGGAACTTCCATGCAAGGCAGTCCA 454

Qy    464 TATCATCTACACTGTCAACCTCTACAGACAGCGGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    455 TGTCATCTACACAGTCAACCTCTACAGCAGTGCCTTCATCCTGGCCTTCATCAGTCTGGA 514

Qy    524 CCGGTACCTGCCATTGTCCACGCCACCAAGTCCTCGGATCCCAGCCCTCCTCCTGGATCCCCAGCCCTCCTCC 583
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    515 CCGCTACCTGCCATCGTCCACGCCACCAAGTCCTCGGATCCCAGCCCTCCTCC 574

Qy    584 AAAGGCAGTCAGTCTATGGGCGTCGGATCCCAGCCCTCCTCCTGACTATACCTGACTTCAT 643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    575 AAAGGTGGTCTATGTGGGTCTCGGATCCCAGCCCTCCTCCTGCTGACTATTCCCGACTTCAT 634

Qy    644 CTTTGCCGACGTCAGCCAGGGGACATCAG 673
          |||||||||||||||||||||||||||||
Db    635 CTTTGCCACGTCAGTCAGTGAGGCAGATGACAG 664


RESULT 15
CD389217
LOCUS       CD389217                       928 bp    mRNA        linear   EST 30-MAY-2003
DEFINITION  AGENCOURT_14302495 NIH_MGC_173 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CD389217
VERSION     CD389217.1   GI:31227843
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 928)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
                         National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
JOURNAL    Unpublished (1999)
COMMENT    Contact: Daniela S. Gerhard, Ph.D.
           Office of Cancer Genomics
           National Cancer Institute / NIH
           Bldg. 31 Rm10A07 Bethesda, MD 20892
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Dr. Jamie Thompson, University of WI
           cDNA Library Preparation: Gina Zastrow-Hayes
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: NDKM74 row:1  column: 08
           High quality sequence start: 8
           High quality sequence stop: 797.
FEATURES             Location/Qualifiers
     source          1..928
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic trophoblasts, made from WA01 stem
                     cells"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_173"
                     /note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
                     LIBR_PRIMING - Oligo dT; METHOD - full-length enriched;
                     LIBR_PROVIDER - Bradfield"
ORIGIN

       Query Match         65.2%;    Score 446.8;  DB 14;   Length 928;
       Best Local Similarity   83.0%;    Pred. No. 1.7e-97;
       Matches 523; Conservative    0; Mismatches 102; Indels    5; Gaps    1;

Qy     44 GCAGGTGCAGGTAGCAGTGACCCTCGAGGCGTTGGTGCTCCGGTAACCACCACGGCTG 103
          ||||||||||||||||||  | |||||| |||||||||||||||| ||||||||||||||
Db     34 GCGGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTG 93

Qy    104 TAGAGCGAGTGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
          ||||||||||||||||||||||||||||||||     |||||||||||||||||||||||
Db     94 GAGAACCAGCGCGGTTACCATGGA----GGGGATCAGTCAGTATATACACTTCAGATAACTACAC 148

Qy    164 TGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAGGACACCCGCTTCCGGGATGAAAA 223
          |||||||||||||||  ||||||||||||  |||||||||||||||||||||||||||||
Db    149 CGAGGAAATGGGGCTCAGGGGACTATGACTCCATGAAGGAACCCGTTTCCGTGAAGAAAA 208

Qy    224 CGTCCATTTCAATAGGATCTTCCTGCCCACCATCTACTTCATCATCATCATCTTAACTGGCAT 283
          |||||||||||||||  |||||||||||||||||||||||  |||||||||||||||||||
Db    209 TGCTAATTTCAATAAAAATTCCTGCCCACCATCTACTTCATCATCATCTTAACTGGCAT 268

Qy    284 AGTCGGCAATGGATTGGTCATCCTGGTCATGGGGTTACCAGAAGAAGTAAGGAGCATGAC 343
          ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    269 TGTGGGCAATGGATTGGTCATCCTGGTCATGGGGTTACCAGAAGAAACTGAGAAGCATGAC 328

Qy    344 GGACAAGTACCGGCTGCACCTGTCAGTGGCTGACTCCTCTTGTCATCATCACACTCCCCTT 403
          |||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db    329 GGACAAGTACAGGCTGCACCTGTCAGTCAGTGGCCGACTCCTCTTGTCATCATCACGGCTTCCCTT 388

Qy    404 CTGGGCAGTTGATGCCATGGCCGGTGCCAAACTCCTACATGTGTACTTTGGGAAATTTTTGTGTAAGGCTGTCCA 463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    389 CTGGGCAGTTGATGCCGTGGCCGGTGCCAAACTGTACTTTGGGAACTTCCATGCAAGGCAGTCCA 448

Qy    464 TATCATCTACACTGTCAACCTCTACAGACAGCGGTTCTCTGATAACTACTC 523
          |||||||||||||||||||||||||||||||||||||||||
Db    449 TGTCATCTACACAGTCAACCTCTACAGCAGTGCCTTCATCCTGGCCTTCATCAGTCTGGA 508

Qy    524 CCGGTACCTGCCATTGTCCATGGGCGTCGGATCCCAGCCCTCCAGCCCTCCTCCTGA 583
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    509 CCGCTACCTGGCCATCGTCCACGCCACCAAGTCCTCGGATCCCAGCCCTCCTCCTGA 568

Qy    584 AAAGGCAGTCATGTGGCGTCTGGATCCCAGCCCTCCTCCTGCTGACTATACCCGACTTCAT 643
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   569 AAAGGTGGTCTATGTTGGCGTCTGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCAT 628

Qy   644 CTTTGCCGACGTCAGCCAGGGGGACATCAG 673
          ||||||| |||||| ||| || ||
Db   629 CTTTGCCAACGTCAGTGAGGCAGATGACAG 658

Search completed: May 17, 2004, 16:09:07
Job time : 1983.17 secs

OM nucleic - nucleic search, using sw model

Copyright GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 17, 2004, 04:56:59 ; Search time 6852.38 Seconds
(without alignments)
10714.982 Million cell updates/sec

Title: US-09-367-052-7
Perfect score: 1694
Sequence: 1 atatacacttctgataacta.........taaaagtcaatgataaactt 1694

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1694	100.0	1877	10	D87747	D87747 Mus musculu
2	1680.4	99.2	1817	10	BC031665	BC031665 Mus muscu
3	1599.8	94.4	175370	2	AC145861	AC145861 Pan trogl
4	1599.8	94.4	185215	2	AC147556	AC147556 Mus muscu
5	1596.2	94.2	1809	10	AB000803	AB000803 Mouse mRN
6	1383.8	81.7	1575	10	MMLESTRPT	X99582 M.musculus
7	1383.8	81.7	3770	10	MMLESTRGN	X99581 M.musculus
8	1168.2	69.0	190149	2	AC122097	AC122097 Rattus no
9	1166.6	68.9	1223	10	MMU59760	U59760 Mus musculu
10	1151.8	68.0	1422	10	MMLCR13	Z80113 Mus musculu
11	1151.8	68.0	1180	10	MMLCR12	Z80112 Mus musculu
12	1139.8	67.3	240931	2	AC107097	AC107097 Rattus no
c 13	1114	65.8	3366	10	MVU65580	U65580 Mus musculu
14	1054.2	62.2	1362	10	RNU54791	U54791 Rattus norv
15	946.4	55.9	1582	6	AX481810	AX481810 Sequence
16	933.4	55.1	1645	9	HSNPYRLA	X71635 H.sapiens m
17	933.4	55.1	1662	9	BC020968	BC020968 Homo sapi
18	933.4	55.1	1664	6	E37853	E37853 Anti-HIV ag
19	933.4	55.1	1664	6	AR222504	AR222504 Sequence
20	933.4	55.1	1664	6	AR380668	AR380668 Sequence
21	933.4	55.1	1664	9	HUMHM89	D10924 Homo sapien
22	933.4	55.1	1664	11	G28514	G28514 SHGC-31461
23	933.4	55.1	1667	9	AK129916	AK129916 Homo sapi
24	933.4	55.1	1679	6	AX287116	AX287116 Sequence
25	933.4	55.1	1679	6	AX548790	AX548790 Sequence
26	933.4	55.1	1679	6	AX705078	AX705078 Sequence
27	933.4	55.1	1737	6	AR015970	AR015970 Sequence
28	933.4	55.1	1737	6	AR060748	AR060748 Sequence
29	933.4	55.1	1737	6	AR070433	AR070433 Sequence
30	933.4	55.1	1737	6	AR103430	AR103430 Sequence
31	933.4	55.1	1737	6	I13753	I13753 Sequence 4
32	933.4	55.1	1737	6	I24455	I24455 Sequence 3
33	933.4	55.1	1944	6	E33806	E33806 Human splic
34	933.4	55.1	1950	9	AF147204	AF147204 Homo sapi
35	933.4	55.1	3733	9	HSCXCR4	Y14739 Homo sapien
36	933.4	55.1	5161	9	AF005058	AF005058 Homo sapi
37	933.4	55.0	172281	9	AC068492	AC068492 Homo sapi
38	931.8	55.0	8747	9	AF052572	AF052572 Homo sapi
39	931.8	55.0	8747	9	HSA224869	AJ224869 Homo sapi
40	931.4	55.0	1670	6	AX409496	AX409496 Sequence
41	931.4	55.0	1670	6	AX663641	AX663641 Sequence
42	931.4	55.0	1670	9	HUMGPCR	L06797 Human (clon
43	908.6	53.6	1050	10	AF452185	AF452185 Rattus no
44	905.4	53.4	1050	10	RNU90610	U90610 Rattus norv
45	903.8	53.4	1637	9	HUMSTSR	M99293 Homo sapien

ALIGNMENTS

RESULT 1
D87747
D87747
LOCUS   D87747   1877 bp   mRNA   linear   ROD 07-FEB-1999
DEFINITION   Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION   D87747
VERSION   D87747.1 GI:1772445
KEYWORDS   murine CXCR-4.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (sites)
AUTHORS   Nagasawa,T., Nakajima,T., Tachibana,K., Iizasa,H., Bleul,C.C.,
               Yoshie,O., Matsushima,K., Yoshida,N., Springer,T.A. and
               Kishimoto,T.

TITLE   Molecular cloning and characterization of a murine pre-B-cell
           growth-stimulating factor/stromal cell-derived factor 1 receptor, a
           murine homolog of the human immunodeficiency virus 1 entry
           coreceptor fusin
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)
MEDLINE   97121456
PUBMED   8962122
REFERENCE 2 (bases 1 to 1877)
AUTHORS   Nagasawa,T.T.
TITLE   Molecular cloning of murine PBSF/SDF-1 receptor
JOURNAL   Unpublished
REFERENCE 3 (bases 1 to 1877)
AUTHORS   Nagasawa,T.
TITLE   Direct Submission
JOURNAL   Submitted (05-SEP-1996) Takashi Nagasawa, Research Institute, Osaka
           Medical Center, Department of Immunology; Murodocho 840, Izumi,
           Osaka 590-02, Japan (E-mail:immunol@osk.threewebnet.or.jp,
           Tel:0725-56-1220, Fax:0725-57-3021)
FEATURES   Location/Qualifiers
     source   1..1877
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="DW34"
                     /cell_type="pre-B cell"
                     /tissue_type="bone marrow"
     gene   1..1877
                     /gene="murine CXCR-4"
     CDS   120..1199
                     /gene="murine CXCR-4"
                     /function="PBSF/SDF-1 receptor"
                     /codon_start=1
                     /evidence=experimental
                     /product="murine CXCR-4"
                     /protein_id="BAA13451.1"
                     /db_xref="GI:1772446"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIIFLTGIVGNGLVILVLVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                     GIVLSCYCIIISKLSHSKGHQKRKALRTTVILILAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"
ORIGIN

Query Match   100.0%; Score 1694; DB 10; Length 1877;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 200

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 260

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   261 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 320

Qy   181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   321 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 380

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGACATGGCAATGGCTGACTGGTACTTTGGG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   381 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGACATGGCAATGGCTGACTGGTACTTTGGG 440

Qy   301 AAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   441 AAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 500

Qy   361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCCATTGTCCACGCCACCAACAGTCAA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   501 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCCATTGTCCACGCCACCAACAGTCAA 560

Qy   421 AGGCCAAGAAGAACTGCTGGCCAGTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   561 AGGCCAAGAAGAACTGCTGGCCAGTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 620

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   621 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 680

Qy   541 GATGACAGGTACATCTGTGACCGGCTTTACCCGGACGTGTGTGATGGTGGTGTTCAA 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   681 GATGACAGGTACATCTGTGACCGGCTTTACCCGGACGTGTGTGATGGTGGTGTTCAA 740

Qy   601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTCGTTACTGC 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   741 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTCGTTACTGC 800

Qy   661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAGGCGCAAGGCCCTCAAGACG 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   801 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAGGCGCAAGGCCCTCAAGACG 860

Qy   721 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   861 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGC 920

Qy   781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 980

Qy   841 CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   981 CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC 1040

Qy   901 CTCTATGCCTTCCTCGGGCCAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCATG 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 CTCTATGCCTTCCTCGGGCCAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCATG 1100

Qy   961 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCGTG 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCGTC 1160

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1220

Qy   1081 AATATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221 AATATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG 1280

Qy   1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1340

Qy   1201 TCTAGTTTTTGTGAGGTTGACTAATTTATATAAATTATATAAATATTGTTTTTTGTTTGTTCATGT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341 TCTAGTTTTTGTGAGGTTGACTAATTTATATAAATTATATAAATATTGTTTTTTGTTTGTTCATGT 1400

Qy   1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTATCGTGTGTAG 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1401 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTATCGTGTGTAG 1460

Qy   1321 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCAGAATGTGTGGTAAATTGAATAAA 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCAGAATGTGTGGTAAATTGAATAAA 1520

Qy   1381 GCTAGCCGGTGATCCTCAGCTGTTGCTCGCATAATCTTCTGCTCAAGCGAGCACCCCACCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 GCTAGCCGGTGATCCTCAGCTGTTGCTCGCATAATCTTCTGCTCAAGCGAGCACCCCACCC 1580

Qy   1441 CCACCCCCACCCCCATTCTTAAATTGTTGGTTATGCTGTGTGATGGTTTGTTTTG 1500

Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 61 Row: d Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES   Location/Qualifiers
     source   1..1817
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:36266 IMAGE:3592479"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene   1..1817
                     /gene="Cxcr4"
                     /note="synonyms: LESTR, fusin, PB-CKR, PBSF/SDF-1, CD184"
                     /db_xref="LocusID:12767"
                     /db_xref="MGI:109563"
     CDS   78..1157
                     /codon_start=1
                     /product="chemokine (C-X-C motif) receptor 4"
                     /protein_id="AAH31665.1"
                     /db_xref="GI:21618782"
                     /db_xref="LocusID:12767"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIFLTGIVGNGLVIVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLILP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"
     misc_feature   246..1004
                     /note="7tm_1; Region: 7 transmembrane receptor (rhodopsin
                     family)"
                     /db_xref="CDD:pfam00001"
ORIGIN

Query Match   99.2%;   Score 1680.4;   DB 10;   Length 1817;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches 1692;   Conservative   0;   Mismatches   0;   Indels   1;   Gaps   1;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   99 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 158

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGACATCTTCCTGCCACCATCTAC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   159 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGACATCTTCCTGCCACCATCTAC 218

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCGTGGCTCATGGGTTAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   219 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCGTGGCTCATGGGTTAC 278

Qy   181 CAGAAGAAGCTAAGGAGCATGACAGACAAGTACCGGCTACACCTGTCAGTGGCTGACCTC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   279 CAGAAGAAGCTAAGGAGCATGACAGACAAGTACCGGCTACACCTGTCAGTGGCTGACCTC 338

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGTACTTTGGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   339 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGTACTTTGGG 398

Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGGCGTTCTC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   1581 CCACCCCACCCCACCCCCATTCTAAATTGTTTGGTTATGCTGTGTGATGGTTGTTTG 1640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501 GTTTTTTTTTGTGTTGTTGTTGTTGGTTTTCTGTAAAAGATGGCACTTAAAAACCAAA 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 GTTTTTTTTTGTGTTGTTGTTGTTGTTTTCTGTAAAAGATGGCACTTAAAAACCAAA 1700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1561 GCCTGAAATGGTGGTAGAAATGCTGGGGTTTTTTGTGTTTTTGTTTTTTTCAGTTTTCA 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701 GCCTGAAATGGTGGTAGAAATGCTGGGGTTTTTTGTGTTTTTGTTTTTTTCAGTTTTCA 1760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621 AGAGTAGATTGACTTCAGTCCCTACAAAATGTACAGTCTTGTATTACATTGTAATAAAAG 1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761 AGAGTAGATTGACTTCAGTCCCTACAAAATGTACAGTCTTGTATTACATTGTAATAAAAG 1820
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1681 TCAATGATAAACTT 1694
           ||||||||||||||
Db   1821 TCAATGATAAACTT 1834

RESULT 2
BC031665
LOCUS   BC031665   1817 bp   mRNA   linear   ROD 07-OCT-2003
DEFINITION Mus musculus chemokine (C-X-C motif) receptor 4, mRNA (cDNA clone
            MGC:36266 IMAGE:3592479), complete cds.
ACCESSION   BC031665
VERSION   BC031665.1 GI:21618781
KEYWORDS   MGC.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1817)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnecch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 1817)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
  COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/

```
Qy   1441 CCACCCCACCCCACCCCATTCTTAAATTGTTGGTTATGCTGTGTGATGGTTGTTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1539 CCACCCCACCCCACCCCATTCTTAAATTGTTGGTTATGCTGTGTGATGGTTGTTG 1598

Qy   1501 GTTTTTTTGTTGTTGTTGTGTTTTTTTTTCTGTAAAAGATGGCACTTAAACCAA 1560
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1599 GTTTTTTTGTTGTTGTTGTTG-TTTTTTTTCTGTAAAAGATGGCACTTAAACCAAA 1657

Qy   1561 GCCTGAAATGGTGGTAGAAATGCTGGGGTTTTTTGTTGTTTTTTTCAGTTTTCA 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1658 GCCTGAAATGGTGGTAGAAATGCTGGGGTTTTTTGTTGTTTTTTTCAGTTTTCA 1717

Qy   1621 AGAGTAGATTGACTTCAGTCCTACAAAATGTACAGTCTTGTATTACATTGTTAATAAAAG 1680
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1718 AGAGTAGATTGACTTCAGTCCTACAAATGTACAGTCTTGTATTACATTGTTAATAAAAG 1777

Qy   1681 TCAATGATAAACTT 1694
          ||||||||||||||
Db   1778 TCAATGATAAACTT 1791


RESULT 3
AC145861
LOCUS       AC145861              175370 bp    DNA     linear   HTG 01-AUG-2003
DEFINITION  Pan troglodytes chromosome UNK clone RP43-21D9, *** SEQUENCING IN
            PROGRESS ***, 39 unordered pieces.
ACCESSION   AC145861
VERSION     AC145861.1  GI:33386791
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 175370)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 175370)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT     --------------- Genome Center ---------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            --------------- Project Information ---------------
            Center project name: C_PT021D09
            ----------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 39 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1     5345: contig of 5345 bp in length
            *       5346     5445: gap of unknown length
            *       5446    17195: contig of 11750 bp in length
            *      17196    17295: gap of unknown length
            *      17296    33250: contig of 15955 bp in length
            *      33251    33350: gap of unknown length
            *      33351    45645: contig of 12296 bp in length
            *      45646    45746: gap of unknown length
            *      45747    61991: contig of 16246 bp in length
            *      61992    62092: gap of unknown length
            *      62093    63423: contig of 1331 bp in length
            *      63424    63523: gap of unknown length
            *      63524    64666: contig of 1143 bp in length


Qy    399 AAATTTTGTGTAAGGCTGTCCATATCATCTACACATGTCAACCTCTACAGCAGCGTTCTC 458

Db    361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCACGGCCACCAACGTCAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    459 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCACGGCCACCAACGTCAA 518

Db    421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    519 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 578

Db    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    579 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 638

Db    541 GATGACAGGTACATCTGTGACCGGCCCTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    639 GATGACAGGTACATCTGTGACCGGCCCTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA 698

Db    601 TTCCAGCATATAATGGTGGGGTCTCTATCATCCTCGCCGGCATCGTCATCCTCTCCGTTACTGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    699 TTCCAGCATATAATGGTGGGGTCTCTATCATCCTCGCCGGCATCGTCATCCTCTCCGTTACTGC 758

Qy    661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGGCCACCAGAAGCGCAAGGCCCTCAAGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    759 ATCATCATCTCTAAGCTGTCACACTCCAAGGGGCCACCAGAAGCGCAAGGCCCTCAAGACG 818

Qy    721 ACAGTCATCCTCATCCTCAGCTTTCTTGCCTGCTGGGCTGCCATATTATGTGGGGATCAGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    819 ACAGTCATCCTCATCCTCAGCTTTCTTGCCTGCTGGGCTGCCATATTATGTGGGGATCAGC 878

Qy    781 ATCGACTCCTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
          ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    879 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 938

Qy    841 CACAAGTGGATCCATCACAGAGGCCCTGCCTTCCACTGTTGCCTGAACCCCATC 900
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    939 CACAAGTGGATCCCATCACAGAGGCCCTGCCTTCCACTGTTGCCTGAACCCCATC 998

Qy    901 CTCTATGCCTTCCTCGGGGGCCAAGTTCAAAAGCTGCTGCCCAGCATGCATGCACTCCATG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    999 CTCTATGCCTTCCTCGGGGGCCAAGTTCAAAAGCTGCTGCCCAGCATGCATGCACTCCATG 1058

Qy    961 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1059 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1118

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1119 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1178

Qy   1081 AATATATATATATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAGAGAG 1140
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1179 AATATATATATATATGATAAAGAACCTTTTTATGTTACACATTTTCCAGATATAGAGAG 1238

Qy   1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1239 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1298

Qy   1201 TCTAGTTTTTGTGAGGTTTGACTTAATTTATATATAAAATATTGTTTTTTTGTTTTGTTCATGT 1260
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   1299 TCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAAATATTGTTTTTTTGTTTTGTTCATGT 1358

Qy   1261 GAATGAGCGGTCTAGGCAGGACCCTGTGGCCAAGTTCTTAGTAGCTGTTTATCGTGTGTAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1359 GAATGAGCGGTCTAGGCAGGACCCTGTGGCCAAGTTCTTAGTAGCTGTTTATCGTGTGTAG 1418

Qy   1321 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAAATGTGTGGTAAATTGAATAAA 1380
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   1419 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAAATGTGGTGGTAAATTGAATAAA 1478

Qy   1381 GCTAGCCGTGATCCGTGCTGTGCTCAGCTGTTGCTCATTCCGAGGAGCACCCCACCC 1440
          |||||||||||||||||||||||||||||||||||||||         |||||||||||||||
Db   1479 GCTAGCCGTGATCCGTGCTGTGCTCAGCTGTTGCTCATAATCTTCATTCCGAGGAGCACCCCACCC 1538
```

```
FEATURES             Location/Qualifiers
     source          1..175370
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="UNK"
                     /clone="RP43-21D9"
     misc_feature    1..5345
                     /note="assembly_name:Contig100"
     misc_feature    5446..17195
                     /note="assembly_name:Contig101"
     misc_feature    17296..33250
                     /note="assembly_name:Contig102"
     misc_feature    33351..45646
                     /note="assembly_name:Contig103"
     misc_feature    45747..61992
                     /note="assembly_name:Contig104"
     misc_feature    62093..63423
                     /note="assembly_name:Contig16"
     misc_feature    63524..64666
                     /note="assembly_name:Contig3"
*    64667   64766: gap of unknown length
*    64767   65994: Contig of 1228 bp in length
     misc_feature    64767..65994
                     /note="assembly_name:Contig45"
*    65995   66094: gap of unknown length
*    66095   67430: contig of 1336 bp in length
     misc_feature    66095..67430
                     /note="assembly_name:Contig48"
*    67431   67530: gap of unknown length
*    67531   68982: contig of 1452 bp in length
     misc_feature    67531..68982
                     /note="assembly_name:Contig5"
*    68983   69082: gap of unknown length
*    69083   70891: contig of 1809 bp in length
     misc_feature    69083..70891
                     /note="assembly_name:Contig56"
*    70892   70991: gap of unknown length
*    70992   72172: contig of 1181 bp in length
     misc_feature    70992..72172
                     /note="assembly_name:Contig60"
*    72173   72272: gap of unknown length
*    72273   73614: contig of 1342 bp in length
     misc_feature    72273..73614
                     /note="assembly_name:Contig67"
*    73615   73714: gap of unknown length
*    73715   75185: contig of 1471 bp in length
     misc_feature    73715..75185
                     /note="assembly_name:Contig68"
*    75186   75285: gap of unknown length
*    75286   76854: contig of 1569 bp in length
     misc_feature    75286..76854
                     /note="assembly_name:Contig69"
*    76855   76954: gap of unknown length
*    76955   78413: contig of 1459 bp in length
     misc_feature    76955..78413
                     /note="assembly_name:Contig73"
*    78414   78513: gap of unknown length
*    78514   79964: contig of 1451 bp in length
     misc_feature    78514..79964
                     /note="assembly_name:Contig75"
*    79965   80064: gap of unknown length
*    80065   81745: contig of 1681 bp in length
     misc_feature    80065..81745
                     /note="assembly_name:Contig77"
*    81746   81845: gap of unknown length
*    81846   83417: contig of 1572 bp in length
     misc_feature    81846..83417
                     /note="assembly_name:Contig79"
*    83418   83517: gap of unknown length
*    83518   84910: contig of 1393 bp in length
     misc_feature    83518..84910
                     /note="assembly_name:Contig80"
*    84911   85010: gap of unknown length
*    85011   86788: contig of 1778 bp in length
     misc_feature    85011..86788
                     /note="assembly_name:Contig81"
*    86789   86888: gap of unknown length
*    86889   88892: contig of 2004 bp in length
     misc_feature    86889..88892
                     /note="assembly_name:Contig82"
*    88893   88992: gap of unknown length
*    88993   90670: contig of 1678 bp in length
     misc_feature    88993..90670
                     /note="assembly_name:Contig83"
*    90671   90770: gap of unknown length
*    90771   93207: contig of 2437 bp in length
     misc_feature    90771..93207
                     /note="assembly_name:Contig84"
*    93208   93307: gap of unknown length
*    93308   96658: contig of 3351 bp in length
     misc_feature    93308..96658
                     /note="assembly_name:Contig85"
*    96659   96758: gap of unknown length
*    96759  100576: contig of 3818 bp in length
     misc_feature    96759..100576
                     /note="assembly_name:Contig86"
*   100577  100676: gap of unknown length
*   100677  104988: contig of 4312 bp in length
     misc_feature    100677..104988
                     /note="assembly_name:Contig87"
*   104989  105088: gap of unknown length
*   105089  108893: contig of 3805 bp in length
     misc_feature    105089..108893
                     /note="assembly_name:Contig88"
*   108894  108993: gap of unknown length
*   108994  112632: contig of 3639 bp in length
     misc_feature    108994..112632
                     /note="assembly_name:Contig89"
*   112633  112732: gap of unknown length
*   112733  116328: contig of 3596 bp in length
     misc_feature    112733..116328
                     /note="assembly_name:Contig90"
*   116329  116428: gap of unknown length
*   116429  120980: contig of 4552 bp in length
     misc_feature    116429..120980
                     /note="assembly_name:Contig91"
*   120981  121080: gap of unknown length
*   121081  126821: contig of 5741 bp in length
     misc_feature    121081..126821
                     /note="assembly_name:Contig92"
*   126822  126921: gap of unknown length
*   126922  131331: contig of 4410 bp in length
     misc_feature    126922..131331
                     /note="assembly_name:Contig93"
*   131332  131431: gap of unknown length
*   131432  138705: contig of 7274 bp in length
     misc_feature    131432..138705
                     /note="assembly_name:Contig94"
*   138706  138805: gap of unknown length
*   138806  145169: contig of 6364 bp in length
     misc_feature    138806..145169
                     /note="assembly_name:Contig95"
*   145170  145269: gap of unknown length
*   145270  150696: contig of 5427 bp in length
     misc_feature    145270..150696
                     /note="assembly_name:Contig96"
*   150697  150796: gap of unknown length
*   150797  159988: contig of 9192 bp in length
     misc_feature    150797..159988
                     /note="assembly_name:Contig97"
*   159989  160088: gap of unknown length
*   160089  167244: contig of 7156 bp in length
     misc_feature    160089..167244
                     /note="assembly_name:Contig97"
*   167245  167344: gap of unknown length
*   167345  175370: contig of 8026 bp in length.
```

```
misc_feature    /note="assembly_name:Contig98"
                167345..175370
ORIGIN          /note="assembly_name:Contig99"

Query Match     94.4%;    Score 1599.8;   DB 2;   Length 175370;
Best Local Similarity  97.8%;   Pred. No. 0;
Matches 1668;  Conservative  0;  Mismatches  22;  Indels  15;  Gaps  4;

Qy      1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGTCTGAGACATGACTCCAACAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 167982 ATATACACTTCTGATAACTACTCTGAGAAGTGGGTCTGAGACATGACTCCAACAAG 168041

Qy     61 GAACCCTGCTTCCGGATGAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168042 GAACCCTGCTTCCGGATGAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC 168101

Qy    121 TTCATCATCTTCTGACTGGCAGTCGGCAATGAGTGGTGATCCGGTCATGGGTTAC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168102 TTCATCATCTTCTGACTGGCAGTCGGCAATGAGTGGTGATCCGGTCATGGGTTAC 168161

Qy    181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168162 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 168221

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGACTGGTACTTTGGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168222 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGACTGGTACTTTGGG 168281

Qy    301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGACAGCGTTCTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168282 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGACAGCGTTCTC 168341

Qy    361 ATCCTGGCCTTCATCAGCCTGACCTGGACCGGTACCGCTCGCCATTGTCCACGCCACCAACAGTCAA 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168342 ATCCTGGCCTTCATCAGCCTGACCTGGACCGGTACCGCTCGCCATTGTCCACGCCACCAACAGTCAG 168401

Qy    421 AGGCCAAGGAAACTGCTGGCTGGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168402 AGGCCAAGGAAACTGCTGGCTGGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 168461

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGAGCGTCAGCAGGGGACATCAGTCAGGGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168462 CTCCTGACTATACCTGACTTCATCTTTGCCGAGCGTCAGCAGGGGACATCAGTCAGGGG 168521

Qy    541 GATGACAGGTACATCTGACCGCCTTTACCCGATAGCCTGTGGATGGTGGTGTTTCAA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168522 GATGACAGGTACATCTGACCGCCTTTACCCGATAGCCTGTGGATGGTGGTGTTTCAA 168581

Qy    601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCGGCATCGTCATCCTCTCGTGTACTGC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168582 TTCCAGCATATAATGGTGGGTCTCATCCTGCCGGCATCGTCATCCTCTCGTGTACTGC 168641

Qy    661 ATCATCATCCTGAAGCTGTCACACTCCAAGGCGCCACCAGAAGCGCAAGGCCCTCAAGACG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168642 ATCATCATCCTGAAGCTGTCACACTCCAAGGCGCCACCAGAGAGCGCAAGGCCCTCAAGACG 168701

Qy    721 ACAGTCATCCTCATCCTAGCTTCTTGCCTGCTGCCATATTATGTGGGATCAGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168702 ACAGTCATCCTCATCCTAGCTTCTTGCCTGCTGCCATATTATGTGGGATCAGC 168761

Qy    781 ATCGACTTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168762 ATCGACTTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATCGTG 168821

Qy    841 CACAAGTGGATCTCCATCACAGAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGACCCCATC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168822 CACAAGTGGATCTCCATCACAGAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGACCCCATC 168881

Qy    901 CTCTATGCCTTCCTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCAATGCCATGCACTCAAACTCCATG 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168882 CTCTATGCCTTCCTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCAATGCCATGCACTCAACTCCATG 168941

Qy    961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGGCGGGTGACACTCTTCCGTC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168942 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGGCGGGTGACACTCTTCCGTC 169001

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCTTATGCAAAGACTTATAT 1080
           |||||||||||||||||||||||||||||||||||||||| |||||||| |||||||
Db 169002 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACACTTATGCAAAGACATATAT 169061

Qy   1081 A----ATATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATA 1136
           |     |||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db 169062 AATATATATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATA 169121

Qy   1137 AGAGACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATG 1196
           |||||||||||||||||||||||||||||||||      |||||||||||||||||||
Db 169122 AGAGACTGACCAGTCTTGTACAGTTTTTTTTTTT---ATTGACTGTTGGGAGTTTATG 169177

Qy   1197 TTCCTCTAGTTTTTGTGAGGTTTGACTTAATTATTATAATA----TTGTTTTTGTTTG 1252
           ||||||||||||||||||||||||||||||||||||||||     ||||||||||||
Db 169178 TTCCTCTAGTTTTTGTGAGGTTTGACTTAATTATTATAAATACTTTTTGTTTTGTTTG 169237

Qy   1253 TTTCATGTGAATGAGCGTCTCAGGCAGGACCTGTCGCCAAGTTCTTAGTAGCTGTTTATCT 1312
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 169238 TTTCATGTGAATGAGCGTCTCAGGCAGGACCTGTCGCCAAGTTCTTAGTAGCTGTTTATCT 169297

Qy   1313 GTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAAATGTGTGTAAAT 1372
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 169298 GTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAAATGTGTGTAAAT 169357

Qy   1373 TGAATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCATAATCCTTCATTCCGAGGAGCA 1432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 169358 TGAATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCATAATCCTTCATTCCGAGGAGCA 169417

Qy   1433 CCCCACCCCACCCCCACCCCCACCCCCATCTTAAATTGTTGGTATGCTGTGTGATGG 1492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 169418 CCCCACCCCACCCCCACCCCCACCCCCATCTTAAATTGTTGGTATGCTGTGTGATGG 169477

Qy   1493 TTTGTTGGTTTTTTTT---GTTGTTGTTGTTTTTGTTTTCTGTAAAAGATGGCAC 1549
           |||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db 169478 TTTGTTGGTTTTTTTTTTTGTTGTTGTTGTTTTTGTTTTCTGTAAAAGATGGCAC 169537

Qy   1550 TTAAAACCAAAGCCTGAAATGTGTAGAAATGCTGGGGTTTTTTTTGTTTTT 1609
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 169538 TTAAAACCAAAGCCTGAAATGTGTAGAAATGCTGGGGTTTTTTTTGTTTTT 169597

Qy   1610 TTCAGTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTGTATTACATT 1669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 169598 TTCAGTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTGTATTACATT 169657

Qy   1670 GTTAATAAAGTCAATGATAAACTT 1694
           ||||||||||||||||||||||||
Db 169658 GTTAATAAAGTCAATGATAAACTT 169682
```

RESULT 4
```
LOCUS       AC147556              185215 bp    DNA     linear   HTG 17-DEC-2003
DEFINITION  Mus musculus chromosome UNK clone RP23-480N14, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC147556
VERSION     AC147556.1  GI:39979584
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 185215)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 185215)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-2003) Genome Sequencing Center, 4444 Forest Park
```

           ------- Summary Statistics -------
           Sequencing vector: M13; 0%
           Sequencing vector: plasmid; 100%
           Chemistry: Dye-primer ET; 0% of reads
           Chemistry: Dye-terminator Big Dye; 100% of reads
           Assembly program: Phrap; version 0.990319
           Consensus quality: 194611 bases at least Q40
           Consensus quality: 196848 bases at least Q30
           Consensus quality: 198294 bases at least Q20
           Insert size: 182000; agarose-fp
           Insert size: 252319; sum-of-contigs
           Quality coverage: 13.47 in Q20 bases; agarose-fp
           Quality coverage: 9.68 in Q20 bases; sum-of-contigs

         * NOTE: This is a 'working draft' sequence. It currently
         * consists of 11 contigs. The true order of the pieces
         * is not known and their order in this sequence record is
         * arbitrary. Gaps between the contigs are represented as
         * runs of N, but the exact sizes of the gaps are unknown.
         * This record will be updated with the finished sequence
         * as soon as it is available and the accession number will
         * be preserved.
         *   1   1119: contig of 1119 bp in length
         *   1120   1219: gap of unknown length
         *   1220   2637: contig of 1418 bp in length
         *   2638   2737: gap of unknown length
         *   2738   5059: contig of 2322 bp in length
         *   5060   5159: gap of unknown length
         *   5160   6203: contig of 1044 bp in length
         *   6204   6303: gap of unknown length
         *   6304   11814: contig of 5511 bp in length
         *   11815   11914: gap of unknown length
         *   11915   18453: contig of 6539 bp in length
         *   18454   18553: gap of unknown length
         *   18554   33684: contig of 15131 bp in length
         *   33685   33784: gap of unknown length
         *   33785   57713: contig of 23929 bp in length
         *   57714   57813: gap of unknown length
         *   57814   84551: contig of 26738 bp in length
         *   84552   84651: gap of unknown length
         *   84652   122285: contig of 37634 bp in length
         *   122286   122385: gap of unknown length
         *   122386   185215: contig of 62830 bp in length.

FEATURES   Location/Qualifiers
     source   1..185215
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP23-480N14"
     misc_feature   1..1119
                     /note="assembly_name:Contig51"
     misc_feature   1220..2637
                     /note="assembly_name:Contig103"
     misc_feature   2738..5059
                     /note="assembly_name:Contig122"
     misc_feature   5160..6203
                     /note="assembly_name:Contig124"
     misc_feature   6304..11814
                     /note="assembly_name:Contig126"
     misc_feature   11915..18453
                     /note="assembly_name:Contig127"
     misc_feature   18554..33684
                     /note="assembly_name:Contig128"
     misc_feature   33785..57713
                     /note="assembly_name:Contig129"
     misc_feature   57814..84551
                     /note="assembly_name:Contig130"
     misc_feature   84652..122285
                     /note="assembly_name:Contig131"
     misc_feature   122386..185215
                     /note="assembly_name:Contig132"

ORIGIN
Query Match   94.4%;   Score 1599.8;   DB 2;   Length 185215;
Best Local Similarity 97.8%;   Pred. No. 0;
Matches 1668;   Conservative   0;   Mismatches   22;   Indels   15;   Gaps   4;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91321 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG   91380

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91381 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC   91440

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91441 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC   91500

Qy   181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91501 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC   91560

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGACTGGTACTTTGGG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91561 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGACTGGTACTTTGGG   91620

Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATCCACTGTCAACCTCTACAGCAGCGTTCTC   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91621 AAATTTTGTGTAAGGCTGTCCATATCATCCACTGTCAACCTCTACAGCAGCGTTCTC   91680

Qy   361 ATCCTGGCCTTCATCAGCCTGACCGGTACCTGCCATTGTCCACGCCACCAACAGTCAA   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db 91681 ATCCTGGCCTTCATCAGCCTGACCGGTACCTGCCATTGTCCACGCCACCAACAGTCAG   91740

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGGTCTGGATCCCAGCCCTC   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91741 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGGTCTGGATCCCAGCCCTC   91800

Qy   481 CTCCTGACTATACCTGACTGATCATCTTTGCGACGCTGACTTGTCCACGCCACCAGTCAGGGG   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91801 CTCCTGACTATACCTGACTGATCATCTTTGCGACGCTGCAGCCAGGGGACATCAGTCAGGGG   91860

Qy   541 GATGACAGGTACATCTGTGACCGGCCTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91861 GATGACAGGTACATCTGTGACCGGCCTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA   91920

Qy   601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCCATCGTCATCCTCCTGTTACTGC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91921 TTCCAGCATATAATGGTGGGTCTCGTCCTGCCCGGCCATCGTCATCCTCCTGTTACTGC   91980

Qy   661 ATCATCATCCTCTAAGCTGTCACACTCGTCACACTGTCCAGAGGCCCAAGGCCCTCAAGACG   720
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 91981 ATCATCATCCTCTAAGCTGTCACACTCGTCACACTGTCCAGAGGCCCAAGGCCCTCAAGACG   92040

Qy   721 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCGTGGCTGCCATTATGTGGGGATCAGC   780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92041 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCGTGGCTGCCATTATGTGGGGATCAGC   92100

Qy   781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGGACTTCGAGAGCATTGTG   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92101 ATCGACTCCTTCATCCTTTGGGGGTCATCAAGCAAGGATGGACTTCGAGAGCATCGTG   92160

Qy   841 CACAGTGGATCTCCATCACAGAGGCCCTGCCCTTCTTCCACTGTTGCCTGGAACCCCATC   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

FEATURES   Location/Qualifiers
     source   1..1809
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL-6"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
                     /dev_stage="E17"
     CDS   72..1151
                     /note="fusin"
                     /codon_start=1
                     /product="murine CXCR-4"
                     /protein_id="BAA19187.1"
                     /db_xref="GI:1816446"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYPIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLFVITLPFWAVDAMA
                     DWYGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLMVVFQHIMVGLVLP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
                     IKQCDFEISIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGHSHSSVSTESESSSFHSS"

ORIGIN

Query Match   94.2%; Score 1596.2; DB 10; Length 1809;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1666; Conservative 0; Mismatches 23; Indels 16; Gaps 4;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTCGAGACTATGACTCCAACAAG   60
           ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   93 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTCGAGACTATGACTCCAACAAG 152

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   153 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC 212

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTATCCTGGTCATGGGTTAC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   213 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTATCCTGGTCATGGGTTAC 272

Qy   181 CAGAAGAGCTAAGGAGCATGACGGACAAGTACCGGCTCAGTGGCTGACCTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   273 CAGAAGAGCTAAGGAGCATGACGGACAAGTACCGGCTCAGTGGCTGACCTC 332

Qy   241 CTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGGCATGGCTGACTGGTACTTTGGG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   333 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGGCATGGCTGACTGGTACTTTGGG 392

Qy   301 AAATTTTTGTGTGTAAGGCTGTCCATATCATCCAATCTTCAACCTCACACTGTCAACCTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   393 AAATTTTTGTGTGTAAGGCTGTCCATATCATCCAATCTTCAACCTCACACTGTCAACCTC 452

Qy   361 ATCCTGGCCTTCATCAGCCTGGACCCGTACCTCGCCATTGTCCACGCCACCAACAGTCAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   453 ATCCTGGCCTTCATCAGCCTGGACCCGTACCTCGCTATTGTCCACGCCACCAACAGTCAG 512

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   513 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 572

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   573 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 632

Db 92161 CACAAGTGGATCCCATCACAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC 92220
           |||||||||||||||||||||||||||||||||||||||
Qy   901 CTTATGCCTTCCTCGGGGCCAAGTTCAAAAGTCTGCCCAGCATGCACTCAACTCCATG 960
           |||||||||||||||||||||||||||||||||||||||
Db 92221 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGTCTGCCCAGCATGCACTCAACTCCATG 92280

Qy   961 AGCAGAGCTCCAGCCTCCAAGATCCTTCCAAAGGAAAGCGGGTGACACTCTTCCGTC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92281 AGCAGAGCTCCAGCCTCCAAGATCCTTCCAAAGGAAAGCGGGTGACACTCTTCCGTC 92340

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGTAACCCTTATGCAAAGACTTATAT 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92341 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACACTTATGCAAAGACATATAT 92400

Qy   1081 A----ATATATATATATATGATAAAGAACTTTTTATGTACACATTTTCCAGATATA 1136
           | |||||||||||||||||||||||||||||||||||||||||||||||||
Db 92401 AATATATATATATATATGATAAAGAACTTTTTATGTACACATTTTCCAGATATA 92460

Qy   1137 AGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTTAATTGACTGTGGGAGTTTATG 1196
           ||||||||||||||||||||||||||||||||| ||||||||||||||
Db 92461 AGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTTT---ATTGACTGTGGGAGTTTATG 92516

Qy   1197 TTCCTCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAATA----TTGTTTTTTGTTTG 1252
           |||||||||||||||||||||||||||||||||||||||| |||||||||||
Db 92517 TTCCTCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAATACTTTTTTTTTGTTTGTTTG 92576

Qy   1253 TTTCATGTGAATGAGCGCTAGGCAGGACCTGGGCCAAGTTCTTAGTAGCTGTTTATCT 1312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92577 TTTCATGTGAATGAGTGTCTAGGCAGGACCTGGGCCAAGTTCTTAGTAGCTGTTTATCT 92636

Qy   1313 GTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGGTAAAT 1372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92637 GTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGGTAAAT 92696

Qy   1373 TGAATAAAGCTAGCCGTGATCCTCAGTGTGTGCTGCATAATCTTCATTCCGAGGAGCA 1432
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92697 TGAATAAAGCTAGCCGTGATCCTCAGTGTGTGCTGCATAATCTTCATTCCGAGGAGCA 92756

Qy   1433 CCCACCCCACCCCACCCCACCCCATTCTTAAATTGTTGGTTATGCGTGTGATGG 1492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92757 CCCACCCCACCCCACCCCACCCCATTCTTAAATTGTTGGTTATGCGTGTGATGG 92816

Qy   1493 TTTGTTTGGTTTTTTT---GTTGTTGTTTGTTTTTTTTCTGTAAAAGATGCGCAC 1549
           |||||||||||||| ||||||||||||||||||||||||||||||||
Db 92817 TTTGTTTGGTTTTTTGTTTTGTTGTTGTTTGTTTTTTTTCTGTAAAAGATGCGCAC 92876

Qy   1550 TTAAAACAAAGCCTGAAAATGGTGGTAGAAAATGCTAGGGTTTTTTTTGTTTGTTTT 1609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 92877 TTAAAACAAAGCCTGAAAATGGTGGTAGAAAATGCTGGGGTTTTTTTGTTTGTTTT 92936

Qy   1610 TTCAGTTTTCAAGAGTAGATTGATTGATTCCTACAAATGTACAGTCTTGTATTACATT 1669
           ||||||||||||||||||| |||| ||||||||||||||||||||||||||
Db 92937 TTCAGTTTTCAAGAGTAGATTGATTCACTCCCTACAAATGTACAGTCTTGTATTACATT 92996

Qy   1670 GTTAATAAAGTCAATGATAAACTT 1694
           ||||||||||||||||||||||||
Db 92997 GTTAATAAAGTCAATGATAAACTT 93021

```
Db   1708 TTCAGTTTTCAAGAGTAGATTGATTCACTCCCTACAAATGTACAGTCTTGTATTACATT 1767
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1670 GTTAATAAAGTCAATGATAAACTT 1694
          ||||||||||||||||||||||||
Db   1768 GTTAATAAAGTCAATGATAAACTT 1792


RESULT 6
MMLESTRPT
LOCUS       MMLESTRPT         1575 bp    mRNA    linear   ROD 11-NOV-1996
DEFINITION  M.musculus mRNA for leukocyte-derived seven transmembrane domain
            receptor.
ACCESSION   X99582
VERSION     X99582.1  GI:1666648
KEYWORDS    lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Moepps,B., Frodl,R., Kessler,H. and Gierschik,P.
  TITLE     cDNA cloning and genomic organization of a leukocyte-derived seven
            transmembrane domain receptor (LESTR) from mouse: a murine
            homologue of the human HIV-1 entry cofactor fusin
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1575)
  AUTHORS   Moepps,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,
            Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK      Revised by author 11-NOV-96
COMMENT     On Nov 12, 1996 this sequence version replaced gi:1657351.
FEATURES             Location/Qualifiers
     source          1..1575
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="B6CBA"
                     /db_xref="taxon:10090"
                     /haplotype="diploid"
                     /sex="female"
                     /tissue_type="thymus"
                     /dev_stage="6-8 weeks old"
     gene            1..1575
                     /gene="lestr"
                     90..1169
                     /gene="lestr"
     CDS             /codon_start=1
                     /product="leukocyte-derived seven transmembrane domain
                     receptor"
                     /protein_id="CAA67894.1"
                     /db_xref="GI:1666649"
                     /db_xref="GOA:P70658"
                     /db_xref="SWISS-PROT:P70658"
                     /translation="MEPISVSIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIPLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLLAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLVLP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISTEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESESSSFHSS"
     polyA_signal    1498..1503
                     /gene="lestr"

ORIGIN

Query Match            81.7%;  Score 1383.8;  DB 10;  Length 1575;
Best Local Similarity  98.3%;  Pred. No. 2.8e-293;
Matches 1434;  Conservative 0;  Mismatches 12;  Indels 13;  Gaps 3;

Qy     1 ATATACACTTCTGATAACTACTCTGAGACAAGTGGGTCTGGAGACTATGACTCCAACAAG 60
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   111 ATATTACACTTCTGATAACTACTCTGAAGAGAAGTGGGTTCTGGAGACTATGACTCCAACAAG 170
```

```
Qy   541 GATGACAGGTACATCTGTGACCGCGCCTTTACCCCCGATAGCCTGTGGATGGTGGTGTTCCAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   633 GATGACAGGTACATCTGTGACCGCGCCTTTACCCCCGATAGCCTGTGGATGGTGGTGTTCCAA 692

Qy   601 TTCCAGCATATAATGGTGGGGCTGGTCTCATCGTCTCCTCCTGTTACTGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   693 TTCCAGCATATAATGGTGGGGCTGGTCTCATCGTCTCCTCCTGTTACTGC 752

Qy   661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGGCACCAGAAGGCAAGGCCCTCAAGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   753 ATCATCATCTCTAAGCTGTCACACTCCAAGGGGCACCAGAAGGCAAGGCCCTCAAGACG 812

Qy   721 ACAGTCATCCTCAGCTTTCTTGCCTGCTGGCTGCCATATTGTGGGGATCAGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   813 ACAGTCATCCTCAGCTTTCTTGCCTGCTGGCTGCCATATTGTGGGGATCAGC 872

Qy   781 ATCGACTCCTTCATCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   873 ATCGACTCCTTCATCTTTGGGGGTCATCAAGCAAGGATGTGACTTCGAGAGCATCGTG 932

Qy   841 CACAAGTGGATCTCCATCACAGAGAGCCCTCGCCTTCTTCACTGTTGCCTGAACCCATC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   933 CACAAGTGGGATCTCCATCACAGAGAGCCCTCGCCTTCTTCACTGTTGCCTGAACCCATC 992

Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGATGCACTCAACTCCATG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   993 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGATGCACTCAACTCCATG 1052

Qy   961 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  1053 AGCAGAGGCTCCAGCCTCAAGATCCTTTTCCAAAGGAAAGCGGGGGTGGCACACTCTTCCGTC 1112

Qy  1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
          ||||||||||||||||||||||||||||||                            
Db  1113 TCCACGGAGTCAGAATCCTCCAGTTTCACTTTT-----ATTGACTGTGGGAGTTTATG 1172

Qy  1081 A---ATATATATATATATGATATGATAAAGACTTTTTATGTTACACATTTTCCAGATATA 1136
          |   |||||||||||||||||||||||||||||||||                       
Db  1173 AATATATATATATATATATATGATATGATAAAGACTTTTTTTATGTTACACATTTTCCAGATATA 1232

Qy  1137 AGAGACTGACCAGTTCTGTACAGTTTTTTTTTTTTTTTAATTGACTGTGGGAGTTTATG 1196
          |||||||||||||||||||||||||||                                
Db  1233 AGAGACTGACCAGTCTGTACAGTTTTTTTTTTT-----ATTGACTGTGGGAGTTTATG 1287

Qy  1197 TTCCCTCAGTTTTTGTGAGGTTTGACTTAAT-----TTGTTTTTTTGTTTG 1252
          ||||||||||||||||||||||||||| |     |||||||||||||||||
Db  1288 TTCCCTCAGTTTTTGTGAGGTTTGACTTAATTTATTATAAATACCCTTTTTTTTGTTTTGTTTG 1347

Qy  1253 TTTGTTGAATGAGGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCT 1312
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1348 TTTCATGTGAATGAGTGTCTCAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCT 1407

Qy  1313 GTGTGTAGGACTGTAGAACTGTAGAGGAAGGAACGAAACATTCCAGAATGTGTCGTAAAT 1372
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1408 GTGTGTAGGACTGTAGAACTGTAGAGAGGAAGGAACTGAAACATTCCAGAATGTGTCGTAAAT 1467

Qy  1373 TGAATAAAGCTAGCCGTGATCCTGTGTTGCTGCTATAATCTCTTCCATCTCCGAGGAGCA 1432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1468 TGAATAAAGCTAGCCGTGATCCTGTGTTGCTGCTATAATCTCTTCCATCTCCGAGGAGCA 1527

Qy  1433 CCCCACCCCCACCCCCACCCCCCACCCCCATTCTAAATTGTTTGGTTATGCTGTGTGATGG 1492
          |||||||||||||||||||||||||                                    
Db  1528 CCCCACCCCCACCCCCACCCCCCACCCCCATTTTTAAATTGTTTGGTTATGCTGTGTGATGG 1587

Qy  1493 TTTGTTGGTTTTTTTT---GTTGTTGTGTTGTTTTTTTTTTTCTGTAAAAGATGGCAC 1549
          |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db  1588 TTTGTTTGGTTTTTTTTTTGTTTTTGTGTTTTTTGTTTTTTTTTCTGTAAAAGATGGCAC 1647

Qy  1550 TTAAAACCAAAGCCTGAAATGCTGAGAAATGCTGGGGTTTTTTTTTGTTTGTTTT 1609
          ||||||||||||||||||||||||||||||||||||                   
Db  1648 TTAAAACCAAAGCCTGAAATGCTGAGAAATGCTGGGGGTTTTTTTTTTGTTTGTTTT 1707

Qy  1610 TTCAGTTTTCAAGAGTAGATTGATTGACTTCAGTCCCTACAAATGACTATTACATT 1669
```

```
Qy    61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 230

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGCATTGGTGATCTGGTCATGGGTTAC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   231 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGCATTGGTGATCTGGTCATGGGTTAC 290

Qy   181 CAGAAGAAGTCAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   291 CAGAGAAGAGTCAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 350

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGGCCATGGCCATGGCTGACGTACTTTGGG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   351 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGGCCATGGCCATGGCTGACGTACTTTGGG 410

Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCACAGCAGCGTTCTC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   411 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCACAGCAGCGTTCTC 470

Qy   361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCCAGCCACCAACAGTCAA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   471 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCCAGCCACCAACAGTCAG 530

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCCAGCCCTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   531 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCCAGCCCTC 590

Qy   481 CTCCTGACTATACCTGACTTCATCTTGCCGACGTCAGCCAGGGGGGACATCAGTCAGGGG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   591 CTCCTGACTATACCTGACTTCATCTTGCCGACGTCAGCCAGGGGGGACATCAGTCAGGGG 650

Qy   541 GATGACAGGTACATCGTGTGACCGCCTTACCCCGATAGCCTGTGATGGTGGTGTTCAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   651 GATGACAGGTACATCGTGTGACCGCCTTACCCCGATAGCCTGTGATGGTGGTGTTCAA 710

Qy   601 TTCCAGCATATAAATGGTGGGTCTCATCCTCGCCCGGCGATCGTCATCCTCTCTGTTACTGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   711 TTCCAGCATATAAATGGTGGGCTCGTCCTGCCCGGCGATCGTCATCCTCTCTGTTACTGC 770

Qy   661 ATCATCATCTCTAAGTGTCACACTCCAAGGCCACCAGAAGCCGCAAGGCCCTCAAGACG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   771 ATCATCATCTCTAAGTGTCACACTCCAAGGCCACCAGAAGCCGCAAGGCCCTCAAGACG 830

Qy   721 ACAGTCATCCCCATCCTAGCTTCTTCTTGCCTGCTGGCTGCCCATATTATGTGGGGATCAGC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   831 ACAGTCATCCCCATCCTAGCTTCTTCTTGCCTGCTGGCTGCCCATATTATGTGGGGATCAGC 890

Qy   781 ATCGACACTCCCTCATCCTTTGGGAGTCATCAAGAAGGATGTGACTTCGAGAGCATTGTG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   891 ATCGACACTCCCTCATCCTTTGGGGTCATCAAGAAGGATGTGACTTCGAGAGCATCGTG 950

Qy   841 CACAAGTGGATCTCCATCACAGAGAGCCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   951 CACAAGTGGATCTCCATCACAGAGAGCCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC 1010

Qy   901 CTCTATGCCTTCCTCCTGGGGCCAAGTTCAAAAGCTCTGCCACTCCAACTCCATG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1011 CTCTATGCCTTCCTCCTGCGGGGCCAAGTTCAAAAGCTCTGCCACTCCAACTCCATG 1070

Qy   961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGTGGACACTCTTCCGTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1071 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGTGGACACATCTTCCGTC 1130

Qy  1021 TCCACCGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACCCTATGCAAAGACTTATAT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1131 TCCACCGGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACACTTATGCAAAGACATATAT 1190

Qy  1081 A---ATATATATATATGATATGATAAAGAACTTTTTATGTTACACCATTTTCCAGATATA 1136
         |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1191 AATATATATATATATATGATATGATAAAGAACTTTTTATGTTACACCATTTTCCAGATATA 1250

Qy  1137 AGAGACTGACCAGTCTTGTACAG---TTTTTTTTTTTTAATTGACTGTTGGGAGTT 1192
         |||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db  1251 AGAGACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTTTTATTGACTGTTGGGAGTT 1310

Qy  1193 TATGTTCCCTAGTTTGTGAGGTTTGACTTAGTTTTTTTTTTTTTT-----TTGTTTTTT 1247
         ||||||||||||||||||||||||||||||||||||||||||    |||||||||
Db  1311 TATGTTCCCTAGTTTGTGAGGTTTGACTTAATTTATATAAAATACCTTTTTTTGTTT 1370

Qy  1248 GTTGTTCATGTGAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTT 1307
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1371 GTTGTTCATGTGAATGAATGAGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTT 1430

Qy  1308 TATCTGTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGG 1367
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1431 TATCTGTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGG 1490

Qy  1368 TAAATTGAATAAAGCTAGCCGTGATCCTCAGCGTGCTGCGCATAATCTTCTTCATTCCGAG 1427
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1491 TAAATTGAATAAAGCTAGCCGTGATCCTCAGCGTGCTGCGCATAATCTTCTTCATTCCGAG 1550

Qy  1428 GAGCACCCCACCCCCACCC 1446
         |||||||||||||||||||
Db  1551 GAGCACCCCCACCCCCACCC 1569
```

RESULT 7

MMLESTRGN

LOCUS   MMLESTRGN   3770 bp   DNA   linear   ROD 02-APR-1997

DEFINITION   M.musculus gene encoding leukocyte-derived seven transmembrane
    domain receptor, strain B6.

ACCESSION   X99581

VERSION   X99581.1   GI:1924959

KEYWORDS   lestr gene; leukocyte-derived seven transmembrane domain receptor.

SOURCE   Mus musculus (house mouse)

ORGANISM   Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE   1

AUTHORS   Moepps,B., Frodl,R., Kessler,H. and Gierschik,P.

TITLE   cDNA cloning and genomic organization of a leukocyte-derived seven
    transmembrane domain receptor (LESTR) from mouse: a murine
    homologue of the human HIV-1 entry cofactor fusin

JOURNAL   Unpublished

REFERENCE   2

AUTHORS   Moepps,B.

TITLE   Direct Submission

JOURNAL   Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,
    Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG

REMARK   Revised by [3]

REFERENCE   3   (bases 1 to 3770)

AUTHORS   Moepps,B.

TITLE   Direct Submission

JOURNAL   Submitted (02-APR-1997) B. Moepps, Universitaet Ulm,
    Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG

COMMENT   On Apr 3, 1997 this sequence version replaced gi:1666646.

FEATURES   Location/Qualifiers

    source   1..3770
               /organism="Mus musculus"
               /mol_type="genomic DNA"
               /strain="B6"
               /db_xref="taxon:10090"
               /haplotype="diploid"
               /tissue_type="thymus"
               /dev_stage="6-8 weeks old"

    exon   <1..110
               /number=1

    gene   join(90..110,2306..3364)
               /gene="lestr"

    CDS   join(90..110,2306..3364)
               /gene="lestr"
               /codon_start=1
               /product="leukocyte-derived seven transmembrane domain
               receptor"
               /protein_id="CAA67893.1"

```
          /db_xref="GI:1666647"
          /db_xref="GOA:P70658"
          /db_xref="SWISS-PROT:P70658"
          /translation="MEPISVSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLP
          TIYFIFLTGIVGNGLVLVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
          DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
          GVWIPALLLTIPDFIFADVSQGDISQGDRYICDRLYPDSLWMVVFQFQHIMVGLVLP
          GIVILSCYCIIISKLSHSKGHQKRKALKTTVILILIAFFACWLPYYVGISIDSFILLGV
          IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSAQHALNSMSRGSSLK
                                                                    "
     misc_feature   105..2305
                    /gene="lestr"
                    /note="splice variant/intron"
     intron         111..2305
                    /gene="lestr"
                    /number=1
     exon           2306..>3770
                    /number=2
     polyA_signal   3693..3698
ORIGIN

Query Match         81.7%;  Score 1383.8;  DB 10;  Length 3770;
Best Local Similarity   98.3%;  Pred. No. 3e-293;
Matches 1434;  Conservative  0;  Mismatches 12;  Indels 13;  Gaps  3;

Qy      1 ATATACACTTCTGATAACTACTCTGAGGAAGTGGGGTCTGGAGACTATGGACTCCAACAAG 60
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   2306 ATATACACTTCTGATAACTACTCTGAGGAAGTGGGGTCTCGGAGACTATGGACTCCAACAAG 2365

Qy     61 GAACCCGTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 120
          ||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2366 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 2425

Qy    121 TTCATCATCTTCTGACTGGCATAGTCGGCAATGGATTGGTGATCCGGTCATGGGTTAC 180
          ||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db   2426 TTCATCATCTTCTGACTGGCATAGTCGGCATAGTGGTGGTCATGGGTTAC 2485

Qy    181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCACCTGTCAGTGGCTGACCTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2486 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCACCTGTCAGTGGCTGACCTC 2545

Qy    241 CTCTTTGTCATCACACTCCCTTCTGGGCAGTGCCATGGCTGATGCCATGGCTGAC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2546 CTCTTTGTCATCACACTCCCTTCTGGGCAGTGCCATGGCTGATGCCATGGCTGAC 2605

Qy    301 AAATTTTTGTGTAAGGCTGTCCATATCATCTCACACTGTCAACCTCACAGCAGCGGTTCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2606 AAATTTTTGTGTAAGGCTGTCCATATCATCTCACACTGTCAACCTCACAGCAGCGGTTCTC 2665

Qy    361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2666 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAG 2725

Qy    421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2726 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 2785

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAAGGGGACATCAGTCAGGGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2786 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAAGGGGACATCAGTCAGGGG 2845

Qy    541 GATGACAGGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGATGGTCGTGTTCAA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2846 GATGACAGGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGATGGTCGTGTTCAA 2905

Qy    601 TTCCAGCATATAATGTGGGTCTCATCCTGCCGGCCATCGTCATCCTCCTGTTACTGC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2906 TTCCAGCATATAATGTGGGTCTCATCCTGCCGGCCATCGTCATCCTCCTGTTACTGC 2965

Qy    661 ATCATCATCTCTAAGCTGTCACACTCCAAGGCGCAAGGCCCTCAAGACG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2966 ATCATCATCTCTAAGCTGTCACACTCCAAGGCGCAAGGCCCTCAAGAGACG 3025

Qy    721 ACAGTCATCCTCATCCTAGCTTCTTTGCCTGCTGCTGCTGCCATATTATGTGGGATCAGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3026 ACAGTCATCCTCATCCTAGCTTCTTTGCCTGCTGCTGCTGCCATATTATGTGGGATCAGC 3085

Qy    781 ATCGACTCCTTCATCCTTTGGAGTCATCAAGCAAGGATGGACTTCGAGAGCATTGTG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3086 ATCGACTCCTTCATCCTTTGGGGTCATCAAGCAAGGATGGACTTCGAGAGCATCGTG 3145

Qy    841 CACAAGTGGATCTCATCACAGAGAGCCCTGCCTTCTTCCACTGTGCCTGAACCCCATC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3146 CACAAGTGGATCTCATCACAGAGAGCCCTGCCTTCTTCCACTGTGCCTGAACCCCATC 3205

Qy    901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGTCTGCCCAGCACGCACTCACCTCCATG 960
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   3206 CTCTATGCCTTCCTCCTGGGGCCAAGTTCAAAAGTCTGCCCAGCACGCACTCACCTCCATG 3265

Qy    961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3266 AGCAGAGGCTCCAGCCTCAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 3325

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTAACCCTTATGCAAAGACTTATAT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3326 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTAACACTTATGCAAAGACATATAT 3385

Qy   1081 A---ATATATATATATATATGATAAAGAACTTTTTATGTACACATTTTCCAGATATA 1136
          |   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3386 AATATATATATATATATATGATAAAGAACTTTTTATGTACACATTTTCCAGATATA 3445

Qy   1137 AGAGACTGACCAGTCTTGTACAG---TTTTTTTTTTTTTAATTGACTGTTGGGAGTT 1192
          ||||||||||||||||||||||||   |||||||||||| ||||||||||||||||||
Db   3446 AGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTTTTATTGACTGTTGGGAGTT 3505

Qy   1193 TATGTTCCTCTAGTTTTTGTGAGTTTGACTTAATTAATTTATATAAATA----TTGTTTTTT 1247
          |||||||||||||||||||||||||||||||||||||||          ||||||
Db   3506 TATGTTCCTCTAGTTTTTGTGAGTTTGACTTAATTAATTTATATAAATAAATTACCTTTTTTTTTGTTT 3565

Qy   1248 GTTTGTTCATGTGAATGAGCGTCAGGCAGGACCTGTCAGGCCAAGTTCTTAGTAGCTGTT 1307
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3566 GTTTGTTCATGTGAATGAGCGTCAGGCAGGACCTGTCAGGCCAAGTTCTTAGTAGCTGTT 3625

Qy   1308 TATCTGTGTGTAGGACTGTAGAACTGTGAGAGGAAGAAACTGAACATTCCAGAATGTGTGG 1367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3626 TATCTGTGTGTAGGACTGTGTAGAACTGTGAGAGGAAGAAACTGAACATTCCAGAATGTGTGG 3685

Qy   1368 TAAATTGAATAAAGCTAGCCTGATCCTCAGCTGTGTCGCTCATTCCGAG 1427
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3686 TAAATTGAATAAAGCTAGCCGTGATCCTCAGCTGTGTCGCTCATTCCGAG 3745

Qy   1428 GAGCACCCCACCCCCACCC 1446
          ||||||||||||||||||
Db   3746 GAGCACCCCACCCCCACCC 3764
```

```
RESULT 8
AC122097
LOCUS       AC122097              190149 bp    DNA     linear   HTG 15-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-108A24, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC122097
VERSION     AC122097.4  GI:25008263
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 190149)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen.H., Alsbrooks,S., Amin,A., Anguiano.D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike.D., Barber,M., Barbazuk,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,J., Burrell,K., Calderon,E.,
```

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,R., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindatne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelemeh,O., Okwuonu,G., Olarmpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,M., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Sheety,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhou,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 190149)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (22-MAY-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 190149)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 15, 2002 this sequence version replaced gi:23269619.
           The sequence in this assembly is a combination of BAC based reads
           and whole genome shotgun sequencing reads assembled using Atlas
           (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
           in the feature table below represents a scaffold in the Atlas
           assembly (a 'contig-scaffold'). Within each contig-scaffold,
           individual sequence contigs are ordered and oriented, and separated
           by sized gaps filled with Ns to the estimated size. The sequence
           may extend beyond the ends of the clone and there may be sequence
           contigs within a contig-scaffold that consist entirely of whole
           genome shotgun sequence reads. Both end sequences and whole genome
           shotgun sequence only contigs will be indicated in the feature
           table.
           --------- Genome Center
               Center: Baylor College of Medicine
               Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----------- Project Information
Center project name: GXCR
Center clone name: CH230-108A24
----------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 179135 bases at least Q40
Consensus quality: 181484 bases at least Q30
Consensus quality: 183011 bases at least Q20
Estimated insert size: 187231; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
*   consists of 4 contigs. The true order of the pieces
*   is not known and their order in this sequence record is
*   arbitrary. Gaps between the contigs are represented as
*   runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
*   as soon as it is available and the accession number will
*   be preserved.
*   1   40067: contig of 40067 bp in length
*   40068 40167: gap of unknown length
*   40168 114660: contig of 74493 bp in length
*   114661 114760: gap of unknown length
*   114761 188186: contig of 73426 bp in length
*   188187 188286: gap of unknown length
*   188287 190149: contig of 1863 bp in length.

FEATURES   Location/Qualifiers
     source   1..190149
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-108A24"

ORIGIN

Query Match   69.0%;   Score 1168.2; DB 2;   Length 190149;
Best Local Similarity 84.6%;   Pred. No. 9.1e-246;
Matches 1456;   Conservative 0;   Mismatches 143;   Indels 122;   Gaps   8;

```
Qy      1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG     60
          ||||||||||| ||||||||||| ||||||||| ||||||||||||||||||||||||||
Db  20662 ATATACACTTCGGATAACTACTCCGAAGAAGTAGGGTCTGGAGACTATGACTCCAACAAG  20721

Qy     61 GAACCCTGCTTCCGGATGAGAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC    120
          |||||||||||||||||||||||||      ||| ||||||||||||||||||||||| |
Db  20722 GAACCCTGCTTCCGGATGAGAAACGAAAAACTTCAACAGGATCTTCCTGCCACCATCTAT  20781

Qy    121 TTCATCATCTTCTTGACTGACTGGCATAGTCGGCAATGATTGGTGATCCTCGGTCATGGGTTAC  180
          || |||||||||||||||||||||||||| || |||||| |||| ||||||||||||||||||
Db  20782 TTTATCATCTTCTTGACTGGCATAGTTGGGCATAGTGGTTGGTAATCCTCGGTCATGGGGTTAC  20841

Qy    181 CAGAAGAAGTCAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCGTCAGTGGCTGACCTC  240
          |||||||||| |||||||||||| ||||||||||||||||||||||||       |||||||||
Db  20842 CAGAAGAAGTGAGGAGCATGACAGACAAGTACCGGCTGCACCTGTCCGGCTGACCTC  20901

Qy    241 CTCTTTGTCATCACACTCCCCCTTCTGGGCAGTTGATGCCATGGCTGGTACTTTGGG    300
          |||||||||||||||||||||| |||||||||||     ||||| |||||||| ||||
Db  20902 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGGACGTGGACGTGGACTGGTACTTTGGG  20961

Qy    301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC    360
          ||||||| ||||||| |||| |||||||||||| |||||||| |||||||||| |||||
Db  20962 AAATTTTATGTAAGGCGTGTGTCATATCATCTACACCGTCAACCTTTACAGCAGTGTTCTC  21021

Qy    361 ATCCTGGCCTTCATCAGCCTGGACCCGGTACCTGTCCACGGCCACCACAGTCAA    420
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  21022 ATCCTGGCCTTCATCAGCCTGGACCCGGTACCTGTCCACGGCCACCACAACAGCCAG  21081

Qy    421 AGGCCAAGGAAACTGCTTGGCTGAAAAGGCAGTCTATGTGGGGCGTCTGGATCCCAGCCCTC  480
          ||||| |||||| ||||||||||||||||| |||| ||||||| |||||||||| |||||||
Db  21082 AGGCCGGAGGAAGTGCTTGGCTGAAAAGGCCCGTTCTATGTGGGGTGTCTGGATCCCCGCCCTC  21141
```

Qy 1534 CTGTAAAGATGGCACTAAAACCAAAGCCTGAAATGTGTGTAGAAATGCTGGGGTTTT 1593
          ||||
Db 22142 CTGTAGAAGATGGCACTAAAACCAAAGCCTGAAATGGT-ATAGAAATGCTGG------ 22193

Qy 1594 TTTGTTTGTTTGTTTTTCAGTTTTCAGAGTAGATTGACTTCAGTCCTACAAATGTAC 1653
Db 22194 ------GTTTTTCTTTCAGTTTCAAGGGTGGATTGATTCAGTTACCTACAAATGTAC 22246

Qy 1654 AGTCTTGTATTACATTGTTAATAAAAGTCAATGATGATAAACTT 1694
          ||||||||
Db 22247 AGTCTTGTATTATGTTGTTTAATAAAAGTCAATGGTAAACTT 22287

RESULT 9
MMU59760
LOCUS   MMU59760   1223 bp   mRNA   linear   ROD 30-SEP-2002
DEFINITION   Mus musculus fusin homolog mRNA, complete cds.
ACCESSION   U59760
VERSION   U59760.1 GI:1527134
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1223)
  AUTHORS   Heesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
  TITLE   Cloning of the mouse fusin gene, homologue to a human HIV-1
             co-factor
  JOURNAL   J. Immunol. 157 (12), 5455-5460 (1996)
  MEDLINE   97113334
   PUBMED   8955194
REFERENCE   2 (bases 1 to 1223)
  AUTHORS   Heesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1996) Pathology, Harvard Medical School, 200
             Longwood Avenue, Boston, MA 02115, USA
FEATURES   Location/Qualifiers
     source   1..1223
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129/Sv"
                     /db_xref="taxon:10090"
     CDS   30..1109
                     /note="seven transmembrane spanning receptor;
                     chemokine-receptor like"
                     /codon_start=1
                     /product="fusin homolog"
                     /protein_id="AAB07725.1"
                     /db_xref="GI:1527135"
                     /translation="MEPISVSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIFLLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GVWIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVFQFQHIMVGLILP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILLLAFFACWLPYYVGISDFILLGV
                     IKQGCDFESIVHKWISITEALAFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTESBSSSFHSS"

ORIGIN

Query Match   68.9%;   Score 1166.6; DB 10; Length 1223;
Best Local Similarity 99.7%;   Pred. No. 1.5e-245;
Matches 1169; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGCTCTGGAGACTATGACTCCAACAAG 60
Db   51 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGCTCTGGAGACTATGACTCCAACAAG 110

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCGCCACCATCTAC 120
Db 111 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCGCCACCATCTAC 170

Qy 121 TTCATCATCTTCTTGACTGGCCAATAGTCGGCAATGAATGGTGTGATCCTGGTCATGGGTTAC 180
Db 171 TTCATCATCTTCTTGACTGGCCAATAGTCGGCAATGAATGGTGTGATCCTGGTCATGGGTTAC 230

Qy 481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTTCAGCCAAGGGGGACATCAGTCAGGGGG 540
Db 21142 CTCCTGACTATCCCTGACACATCATCTTCGGCCGATGTCAGCCAGGGGGAC---------- 21189

Qy 541 GATGACAGGTACATCTGTGACCGCCTTTACCCGATAGCCTGTGGATGGTGGTGTTCAA 600
Db 21190 ---GGCAGGTACATCTGTGACCGCCTTTACCCGACAGCCTGTGGATGGTGTGTTCCAG 21246

Qy 601 TTCCAGCATATAATGGTGGTCTCATCCTGCCCGGCATGTCATCCTCCTGTTACTGC 660
Db 21247 TTCCAGCACATCATGATGGTGGGTCTCATCCTGCCCGGGCATCGTCCTCGTGTTACTGC 21306

Qy 661 ATCATCATCTCAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACG 720
Db 21307 ATCATCATCTCCAAGCTGTCACACTCCAAGGGGCCACCAGAAGCGCAAGGCCCTCAAGACT 21366

Qy 721 ACAGTGCATCATCCTCCAGCTTCTTCTTTGCCTGCTGGCTTCCATATTATGTGGGGATCAGC 780
Db 21367 ACGGTCATCTTATCCTGCTTTCTTTGCCTGCTGGCTACCGTATTACGTGGGGATCAGC 21426

Qy 781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
Db 21427 ATCGATTCCTTCATCCTTTGGAGGTCATCAAGCAAGGATGTGAGTTCGAGAGCGTCGTG 21486

Qy 841 CACAAGTGGATCTCCATCACAGAGAGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATC 900
Db 21487 CACAAGTGGATCTCCATCACGCGAGAGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC 21546

Qy 901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCAGCATGCAGCTCAACTCCATG 960
Db 21547 CTCTACGCGCCTTCCTCGGGGCCCAAATTCAAGAGCTCCGCGCAGCATGCAGCTCAATTCCATG 21606

Qy 961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAAGGAAACGGGGGGTGACACTCTTCCGTC 1020
Db 21607 AGCAGAGGCTCCAGCTTCAAGATCCTTTCAAAAGGAAACGGGGTGGCACTCTTCCGTC 21666

Qy 1021 TCCACGGGAGTCAGAATCCTCCAGTTTCACTCCAGTTTTATGCAAAGACTTATAT 1080
Db 21667 TCCACAGAGTCAGAATCCTCCAGTTTCACTCCAGTTTACACTTACCCAAAGACACATATAT 21726

Qy 1081 AATATATATATATAT--------GATAAAGAACTTTTTTTATGTTACACATTTCCAG 1131
Db 21727 AATATATATATATATATATATATATGATGACAAAGAACTTTTTTTATGTTACACATTTCCCAG 21786

Qy 1132 ATATAAGAAGACTGACCAGTCTGTACAGTTTTTTTTTTTTTTAATTGACTGTTGGGAGT 1191
Db 21787 ATATAAAGACTGACCAATCTTGACAGTTTTTTTTTTTTTTATTGACTGTTGGAGTTTTC 21846

Qy 1192 TTATGTTCCTCT------AGTTTTGTGAGGTTGACTTAATTATATAAATAT---- 1239
Db 21847 TCTAGTTTTTTGTGTGTGTGTTGAGGTTGTGTGAGGTTGACTTAATTTATATAAATATTGGTT 21906

Qy 1240 ------TGTTTTTTTGTTTGTTTTGTCTGTTTCACGTGAATGAGTGTTTCACGGTTTATGTTGT 1293
Db 21907 TTTTTTCTGTTTGTTTGTCTGTTTCACGGTGAATGAGTGTTTATGTTTC 21966

Qy 1294 TCTTAGTAGCTGTTTATCTGTGTGTAGGACTGTGAGAACTGTAGAGGAGGAGAAACTGAACAT 1353
Db 21967 TCTTAGTAGCTGTTTATCTGTGTGTAGGACTGTGAGAGACTGTAGAGGAGGAGAAACTGAACGC 22026

Qy 1354 TCCAGAATGTGTGTAAATTGAATAAAGCTAGCCGTGATCCTGAGTCCTGTTGCTGCATAT 1413
Db 22027 TCCAGAATGTGTGGTAAATCGGATAAAGCTAGAGTGACCCTGAGTCTGGCTGCATAAT 22086

Qy 1414 CTCTTCATTCCGAGGAGCACCCCACCCCCACCCCCACCCCATTCTTAAATTGTT 1473
Db 22087 CTCTT------------------CGCCGCCGCCACCCCCGTTCCTGAATTGTT 22120

Qy 1474 TGGTTATGCTGTGTGATGGTTTGTTTGGTTTTGTGTTGTTGTTTGTTTTTTTTTT 1533
Db 22121 TGGTTGTGCTGTGTG----------------GTTTCT 22141

```
DEFINITION  Mus musculus lcr-1 gene, exon 2.
ACCESSION   Z80113
VERSION     Z80113.1  GI:1655638
KEYWORDS    CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1422)
  AUTHORS   Schubel,A., Burgstahler,R. and Lipp,M.
  TITLE     The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
            High expression in thymus and lymphoid tissues
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1422)
  AUTHORS   Lipp,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
            Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
            D-13122, GERMANY
FEATURES             Location/Qualifiers
     source          1..1422
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV"
                     /db_xref="taxon:10090"
                     /clone_lib="lambda fix (TM) II"
                     /dev_stage="adult"
     intron          <1..259
                     /number=1
     gene            260..1422
                     /gene="lcr-1"
     exon            260..>1422
                     /gene="lcr-1"
                     /product="CXCR-4"
                     /standard_name="CXC chemokine receptor 4"
                     /number=2
                     /function="receptor for chemokine SDF-1"
ORIGIN

Query Match         68.0%;    Score 1151.8;   DB 10;   Length 1422;
Best Local Similarity   99.4%;   Pred. No. 2.7e-242;
Matches 1156;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;

Qy    1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  260 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG  319

Qy   61 GAACCCTGCTTCCGGATGAAAACGTCATTTCAATAGGATCTTCCTGCCACCATCTAC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  320 GAACCCTGCTTCCGGATGAAAACGTCATTTCAATAGGATCTTCCTGCCACCATCTAC  379

Qy  121 TTCATCATCTTGACTGGCATAGTCGGCAATAGTCGGCAATGGATTGGTGATCCTGGTTAC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  380 TTCATCATCTTGACTGGCATAGTCGGCAATAGTCGGCAATGGATTGGTGATCCTGGTTAC  439

Qy  181 CAGAAGAAGCTAAGGAGCATGACGGACCAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  440 CAGAAGAAGCTAAGGAGCATGACGGACCAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  499

Qy  241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGCATGGCCATGGCTGACTGGTACTTTGGG  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  500 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGCATGGCCATGGCTGACTGGTACTTTGGG  559

Qy  301 AAATTTTGTGTAAGGTAAGGCTGTCCATATCATCTCACACTGTCAACCTTCTCACGAGCGTTCTC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  560 AAATTTTGTGTAAGGTAAGGCTGTCCATATCATCTCACACTGTCAACCTTCTCACGAGCGTTCTC  619

Qy  361 ATCCTGGCCTTCATCAGCTGGACCCGGACCTGGACCTGTCCAGCGCCACCAACAGTCAA  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  620 ATCCTGGCCTTCATCAGCTGGACCCGGACCTGGACCTGTCCAGCGCCACCAACAGTCAA  679

Qy  421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGCGTCTGGATCCCAGCCCTC  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  181 CAGAAGAAGCTAAGGAGCATGACGGACCAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  231 CAGAAGAAGCTAAGGAGCATGACGGACCAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  290

Qy  241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGCATGGCCATGGCTGACTGGTACTTTGGG  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  291 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGCATGGCCATGGCTGACTGGTACTTTGGG  350

Qy  301 AAATTTTGTGTAAGGCTGTCCATATCATCTCAACCTCTACAGCAGGCGTTCTC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  351 AAATTTTGTGTAAGGCTGTCCATATCATCTCAACCTCTACAGCAGGCGTTCTC  410

Qy  361 ATCCTGGCCTTCATCAGCTGGACCCGGACCTGTCCAGCGCCACCAACAGTCAA  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  411 ATCCTGGCCTTCATCAGCTGGACCCGGACCTGTCCAGCGCCACCAACAGTCAA  470

Qy  421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGCGTCTGGATCCCAGCCCTC  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  471 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGCGTCTGGATCCCAGCCCTC  530

Qy  481 CTCCTGACTATACCTGACTTCATCTTTGCCACGTCAGCCAGGGGACATCAGTCAGGGG  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  531 CTCCTGACTATACCTGACTTCATCTTTGCCACGTCAGCCAGGGGACATCAGTCAGGGG  590

Qy  541 GATGACAGGTACATCTGACCGCCTTACCCGATAGCCGTGGATGGTGGTGTTCAA  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  591 GATGACAGGTACATCTGACCGCCTTACCCGATAGCCGTGGATGGTGGTGTTCAA  650

Qy  601 TTCCAGCATATAAATGGTGGGTCTCATCCTGGCCGGCATCGTCATCCTCTCCTGTTACTGC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  651 TTCCAGCATATAAATGGTGGGTCTCATCCTGCGCCGGCATCGTCATCCTCTCCTGTTACTGC  710

Qy  661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCGCAAGGCCCTCAAGACG  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  711 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCGCAAGGCCCTCAAGACG  770

Qy  721 ACAGTCATCCTCATCCCTAGCTTTCTTTGCCTGGCTGCCATATTATGTGGGGATCAGC  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  771 ACAGTCATCCTCATCCCTAGCTTTCTTTGCCTGGCTGCCATATTATGTGGGGATCAGC  830

Qy  781 ATCGACTCCTTCATCCTTTTGGGGGGTCATCAAGCAAGGATGTGACTTCGAGCATCGTG  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  831 ATCGACTCCTTCATCCTTTTGGGGGGTCATCAAGCAAGGATGTGACTTCGAGCATCGTG  890

Qy  841 CACAAGTGGATCCATCACAGAGGCCCTGCCCTTCTTCCACTGTGCCTGAACCCCATC  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  891 CACAAGTGGATCCATCACAGAGGCCCTGCCCTTCTTCCACTGTGCCTGAACCCCATC  950

Qy  901 CTCTATGCCTTCTCCTGGGGGCCAAGTTCAAAAGCTCTGCCCCAGCATGCACTCAACTCATG  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  951 CTCTATGCCTTCTCCTGGGGGCCAAGTTCAAAAGCTCTGCCCCAGCATGCACTCAACTCCATG 1010

Qy  961 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1011 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1070

Qy 1021 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1071 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1130

Qy 1081 AATATATATATATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAAGAG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1131 AATATATATATATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAAGAG 1190

Qy 1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTT 1173
         |||||||||||||||||||||||||||||||
Db 1191 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTT 1223
```

RESULT 10
MMLCR13
LOCUS   MMLCR13   1422 bp   DNA   linear   ROD 31-OCT-1996

```
COMMENT      On Nov 21, 1997 this sequence version replaced gi:1542888.
FEATURES             Location/Qualifiers
     source          1..1180
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129Sv"
                     /db_xref="taxon:10090"
                     /tissue_type="thymus"
                     /dev_stage="adult"
     gene            1..1180
                     /gene="lcr-1"
     CDS             1..1074
                     /gene="lcr-1"
                     /function="receptor for chemokine SDF-1"
                     /standard_name="CXC chemokine receptor 4"
                     /codon_start=1
                     /product="CXCR-4"
                     /protein_id="CAB02202.1"
                     /db_xref="GI:1542889"
                     /db_xref="GOA:P70658"
                     /db_xref="SWISS-PROT:P70658"
                     /translation="MEPISIYTSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIFLPTI
                     YFIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADW
                     YFGKFLCKAVHIIYTVNLYSSVLLIAFISLDRYLAIVHATNSQRPRKLLAEKAVVGV
                     WIPALLLTIPDFIFADVSQGDISQGDDRYICDRLYPDSLWMVVFQFQHIMVGLLLPGI
                     VILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYVGISIDSFILLGVIK
                     QGCDFPSIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLKIL
                     SKGKRGGHSSVSTESESSFHSS"
     exon            <1..15
                     /gene="lcr-1"
                     /number=1
     exon            16..>1180
                     /gene="lcr-1"
                     /number=2
ORIGIN

Query Match         67.3%;  Score 1139.8;  DB 10;  Length 1180;
Best Local Similarity   99.2%;  Pred. No. 1.1e-239;
Matches 1156;   Conservative   0;  Mismatches   7;  Indels   2;  Gaps   1;

Qy        1 ATATACACTTCTGATAACTACTTCTGAGAAGTGGGGTCTGGAGACATGACTGACTCCAACAAG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       16 ATATACACTTCTGATAACTACTTCTGAGAAGTGGGGTCTGGAGACATGACTGACTCCAACAAG  75

Qy       61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCATCCATCTAC  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       76 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCATCCATCTAC  135

Qy      121 TTCATCATCTTCTTGACTGGCATAGTCCGGCAATGAGTGGTGATCCTGGTCATGTGGTTAC  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      136 TTCATCATCTTCTTGACTGGCATAGTCCGGCAATGAGTGGTGATCCTGGTCATGTGGTTAC  195

Qy      181 CAGAAGAGCTAAGGAGCATGACGGACAAGCACCGGCTGCACCTGTCAGTGGCTGACCTC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      196 CAGAAGAGCTAAGGAGCATGACGGACAAGCACCGGCTGCACCTGTCAGTGGCTGACCTC  255

Qy      241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGATGCTGACTACTTTGGG  300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      256 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGATGCTGACTACTTTGGG  315

Qy      301 AAATTTTTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      316 AAATTTTTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC  375

Qy      361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCCATTGTCCACGCCACCAACAGTCAA  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      376 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCCATTGTCCACGCCACCAACAGTCAA  435

Qy      421 AGGCCAAGGAAACTGCTGGCTGAAAAAGGCAGTCTATGTGGGCGTCTGAATCCCAGCCCTC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      436 AGGCCAAGGAAACTGCTGGCTGAAAAAGGCAGTCTATGTGGGCGTCTGAATCCCAGCCCTC  495

Qy      481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG  540
```

```
Db      680 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC  739

Qy      481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      740 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG  799

Qy      541 GATGACAGGTACATCGTGACCGGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA  600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      800 GATGACAGGTACATCGTGACCGGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA  859

Qy      601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTGTTACTGC  660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      860 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTGTTACTGC  919

Qy      661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCCAGAAGGCAAGGCCCTCAAGACACG  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      920 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCCAGAAGGCAAGGCCCTCAAGACACG  979

Qy      721 ACAGTCATCCTCCATCCTAGCTTCTTCTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGC  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      980 ACAGTCATCCTCCATCCTAGCTTCTTCTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGC  1039

Qy      781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCCATTGTG  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1040 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCCATTGTG  1099

Qy      841 CACAAGTGGATCTCATCACAGAGCCCCTGCCTTCTTCCACTGTTGCTGAACCCCATC  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1100 CACAAGTGGATCTCATCACAGAGCCCCTGCCTTCTTCCACTGTTGCTGAACCCCATC  1159

Qy      901 CTCTATGCCTTCCTCGGGGCCCAAGTTCAAAAGTCTGCCCAGCATGCACTCAACTCCATG  960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1160 CTCTATGCCTTCCTCGGGGCCCAAGTTCAAAAGTCTGCCCAGCATGCACTCAACTCCATG  1219

Qy      961 AGCAGAGGCCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1220 AGCAGAGGCCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC  1279

Qy     1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1280 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT  1339

Qy     1081 AATATATATATATATGATAAAGACTTTTTATGTTACACATTTTCAGATATAAGAG  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1340 AATATATATATATATGATAAAGACTTTTTTATGTTACACATTTTCCAGATATAAGAG  1399

Qy     1141 ACTGACCAGTCTTGTACAGTTTT  1163
            |||||||||||||||||||||||
Db     1400 ACTGACCAGTCTTGTACAGTTTT  1422
```

```
RESULT 11
MMLCR12
LOCUS       MMLCR12                 1180 bp    mRNA    linear    ROD 18-NOV-1997
DEFINITION  Mus musculus lcr-1 gene.
ACCESSION   Z80112
VERSION     Z80112.1  GI:2632100
KEYWORDS    CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1180)
  AUTHORS   Schubel,A., Burgstahler,R. and Lipp,M.
  TITLE     The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
            High expression in thymus and lymphoid tissues
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1180)
  AUTHORS   Lipp,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
            Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
            D-13122, GERMANY
```

```
Db    496  CTCCTGACTATACCTGACTTCATCTTGCCGACGTGCAGCGGGACATCAGTCAGGGG  555
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541  GATGACAGGTACATCTGTGACCGGCCTTACCCCGATAGCCTGTGATGGTGGTGTTCAA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    556  GATGACAGGTACATCTGTGACCGGCCTTACCCCGATAGCCTGTGATGGTGGTGTTCAA  615

Qy    601  TTCCAGCATATAATGGTGGGTCATCATCCTGCCCGGCATCGTCATCCTCTCGTTACTGC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    616  TTCCAGCATATAATGGTGGGTCATCATCCTGCCCGGCATCGTCATCCTCTCGTTACTGC  675

Qy    661  ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGGCCCTCAAGACG  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    676  ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGGCCCTCAAGACG  735

Qy    721  ACAGTCATCTCCATCCTAGCTTCTTGCCTCGCTGGCCATATTATGTGGGGATCAGC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    736  ACAGTCATCCTCCATCCTTCTTGCCTCGCTGCCATATTATGTGGGGATCAGC  795

Qy    781  ATGGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    796  ATGGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  855

Qy    841  CACAAGTGGATTCCATCACAGAGGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATC  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    856  CACAAGTGGATTCCATCACAGAGGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATC  915

Qy    901  CTCTATGGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    916  CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG  975

Qy    961  AGCAGAGGCCTCCAGCCTCAAGATCCTTCCAAAGGAAAAGGGGGTGGACACTCTTCCGTC  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    976  AGCAGAGGCCTCCAGCCTCAAGATCCTTCCAAAGGAAAAGGGGGTGGACACTCTTCCGTC  1035

Qy    1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1036 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT  1095

Qy    1081 AATATATATATATATATGATAAAGAACTTTTTTATGTTACAC--ATTTTCCAGATATAAG  1138
           ||||||||||||||||||||||||||||||||||||||||||  |||||||||||||||
Db    1096 AATATATATATATATATTGATAAAGACTTTTTTATGTTACACCATTTTTCCAGATATAAG  1155

Qy    1139 AGACTGACCAGTCTTGTACAGTTTT  1163
           |||||||||||||||||||||||||
Db    1156 AGACTGACCAGTCTTGTACAGTTTT  1180
```

RESULT 12
AC107097/c
LOCUS   AC107097   240931 bp   DNA   linear   HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-175O10, *** SEQUENCING IN PROGRESS
            ***.
ACCESSION   AC107097
VERSION   AC107097.5 GI:30580624
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 240931)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandatanaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chu,J.,
            Chacko,J., Chavez,D., Chen,D., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Flagg,N., Forbes,L., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,V., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pai,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,V., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,D., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 240931)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 240931)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On May 13, 2003 this sequence version replaced gi:23115278.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            --------------- Genome Center
               Center: Baylor College of Medicine
               Center code: BCM
               Web site: http://www.hgsc.bcm.tmc.edu/
               Contact: hgsc-help@bcm.tmc.edu
            --------------- Project Information
               Center project name: GKEK
               Center clone name: CH230-175O10

Summary Statistics

```
        ---------- Summary Statistics
        Assembly program: Atlas 3.0;
        Consensus quality: 216693 bases at least Q40
        Consensus quality: 219254 bases at least Q30
        Consensus quality: 221224 bases at least Q20
        Estimated insert size: 228277; sum-of-contigs estimation
        Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
        -------------
    *   NOTE: Estimated insert size may differ from sequence length
    *      (see http://www.hgsc.bcm.tmc.edu/docs/Genbank draft data.html).
    *   NOTE: This is a 'working draft' sequence. It currently
    *   consists of 1 contigs. Gaps between the contigs
    *   are represented as runs of N. The order of the pieces
    *   is believed to be correct as given, however the sizes
    *   of the gaps between them are based on estimates that have
    *   provided by the submittor.
    *   This sequence will be replaced
    *   by the finished sequence as soon as it is available and
    *   the accession number will be preserved.
    *        1 240931: contig of 240931 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..240931
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-175010"
     misc_feature    1..1249
                     /note="wgs end_extension
                     clone_end:Sp6"
     misc_feature    complement(5161..5587)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH357332"
     misc_feature    complement(233579..233220)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH357330"
     misc_feature    235468..237558
                     /note="wgs end_extension
                     clone_end:T7"
     misc_feature    239450..240931
                     /note="wgs end_extension
                     clone_end:T7"
ORIGIN
```

```
Query Match       65.8%;    Score 1114;   DB 2;     Length 240931;
Best Local Similarity  84.2%;  Pred. No. 7.4e-234;
Matches 1400;  Conservative   0;  Mismatches 140;   Indels 122;   Gaps   8;

Qy       60  GGAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTA  119
                 ||||||||||||||||||||||||||||||||| | |||||||||||||||||||||||
Db   237558  GGAACCCCGCTTCCGGGATGAAAACGAAAACTTCAACAGGATCTTCCTGCCCACCATCTA  237499

Qy      120  CTTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGGGTTA  179
                 |||||||||||||||||||||||||||||||||| | |||||||||||||||||||||
Db   237498  TTTTATCATCTTCTTGACTGGCATAGTGGGCATAGTGGGGCAATGGGTTGGTGAATCCTGGTCATGGGTTA  237439

Qy      180  CCAGAAGAAGCTAAGGAGCATGACGGACAAGCAAGTACCCGCTCGACCCTGTCAGTGGCTGACCT  239
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237438  CCAGAAGAAGCTGAGGAGCATGACGGACAGCAAGTACCCGCTCGACCTGTCCGTGGGCTGACCT  237379

Qy      240  CCTCTTTGTCATCACACTCCCCTTCTCGGCAGTTGATGCCATGGCTGACTTGGTTACTTTGG  299
                 ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237378  CCTCTTTGTCATCACACTCCCCCTTCTCGGCAGTTGATGCCATGGCTGACTTGGTTACTTTGG  237319

Qy      300  GAAATTTTTGTGTAAGGCTGTCCCATATCATTCATCTCACACCTTCAACCTCTACAGCAGCGTTCT  359
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237318  GAAATTTTTATGTAAGGCTGTCCCATATCATCCATCTCACACCTTCAACCTTTACAGCAGTGTTCT  237259

Qy      360  CATCCTCCGGCCTTCATCAGCTGCTTCCCCGACCTGACCGGGTACCTCCGCCACCAACAGTCA  419
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   237318  CATCCTCGGCCTTCATCAGCTGCTTCCCCGACCTGACCGGGTACCTCCGCCACCAACAGTCA  236160
```

Sequence alignment (top section)

```
Db   237258  CATCCTGGCCTTCATCAGCCTGGACCGCTACCTTGCCATTGTTCCAGCGCCACCAACAGCCA  237199
Qy      420  AAGGCCAAGGAAACTGCTGGCTGAAAACGCAGTCATGTGGGCGTCTGGATCCCCAGCCCT  479
Db   237198  GAGGCCGAGGAGGCTGCTGGCTGGCTGAAAAAGGCCGTGTCTATGTGGGTGTCTGGATCCCGTCCCGCCT  237139
Qy      480  CCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGG  539
Db   237138  CCTCCTGACTATCCCTGACATCATCTTCGCGATGTCAGCCAGGGGAC----------  237090
Qy      540  GGATGACAGGTACATCTGTGACCGGCCTTACCCTGACATAGCCTGTGGATGTGTGGTTTCA  599
Db   237089  ----GGCAGGTACAATCTGTGACCCGGCCCTTTACCCCGACAGCCCTGTGGATGGTGGTGTTCCA  237034
Qy      600  ATTCCAGCACATATAATGGTCATCTCCTGAGTCATCAAGCAACGGATGTGACTTCGAGAGCATTGT  659
Db   237033  GTTCCAGCACATCATGGGTGGGTCATCCTGAGTCATCAAGCAACGGATGTGACTTCCTCTGTTACTG  236974
Qy      660  CATCATCATCTCTAAGCTGTCACACTCCAAGGCCACCAGAGCCCACCAGAGGCGCAAGGCCCTCAAGAC  719
Db   236973  CATCATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAGCGCAAGGCCCTCAAGAC  236914
Qy      720  GACAGTCATCCTCCTCTCATCCCTAGCTTCTTTGCCTGCTGGCTGCCATATATGTGGGGATCAG  779
Db   236913  TACGGTCATCTTCATCCCTTATCCTTGGCTTTCTTTCCTTGCCTGGCTGCCACGTATTACGTGGGGATCAG  236854
Qy      780  CATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAACGGATGTGACTTCGAGAGCATTGT  839
Db   236853  CATCGATTCCTTCATCCTTTTGGAGGTCATCAACAAGGATGTGAGTTCGAGAGCGTCGT  236794
Qy      840  GCACAAGTGGATCCCATCACAGAGGGCCCTCGCCTTCTCCCACTGTTGCCTGCTGACCCCAT  899
Db   236793  GCACAAGTGGATCTCCATCACGGAGGGCCCTCGCCTTCTTCCACTGTTGCCTGCTGACCCCCAT  236734
Qy      900  CCTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTGACTCAACTCCAT  959
Db   236733  CCTCTACGCCTTCCTCGGGGGCCAAATTCAAGAGCTCCGGCAGCCACTGACTCACTCAATTCCAT  236674
Qy      960  GAGCAGAGAGGCTCAGCCTCAGATCCTTCAAAGATCCTTCCAAAGGAAACGGGGTGACACTCTTCCGT  1019
Db   236673  GAGCAGAGAGGCTCAGCCTCAGATCCTTCAAGATCCTTCCAAAGGAAACGGGGTGACACTCTTCCGT  236614
Qy     1020  CTCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTCGCAAAGACTTATA  1079
Db   236613  CTCCACAGAGTCAGAATCCTCAGTTTCACTCCAGCTCGCAAAGACATATA  236554
Qy     1080  TAATATATATATATAT--------GATAAGAACTTTTTTATGTTACACATTTTCCA  1130
Db   236553  TAATATATATATATATATATATATATATATGACAAAGAACTTTTTTTATGTTACACATTTCCCA  236494
Qy     1131  GATATAAGAGACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTAATTGACTGTTGGGAG  1190
Db   236493  GATATAAAGACTGACCAATCTGTACAGTTTTTTTTTTTATTGACTGTTGGAGTTTATGTTT----  236434
Qy     1191  TTTATGTTCCTCT-------AGTTTTTGTGAGGTTTGACTTAATTATATAAATAT--  1239
Db   236433  CTCTAGTTTTTGTGTGTGTGTGTTTGACTTGACTTAATTAATTAATTAAAAATATTGGT  236374
Qy     1240  -----TGTTTTTTGTTTGTTTCATGTGAATGAGCGTCTAGGCAGGACCGTGGCCAAG  1292
Db   236373  TTTTTTCGTTGTTTGTCGTTTCACGTGAATGAGTGTCTAGGCAGGACCGTCGGCCAAG  236314
Qy     1293  TTCTTAGTAGCTGTTTATCTGTGTAGGACTGTGAGGACTGTAGAGGAGAAACTGAACA  1352
Db   236313  TTCTTAGTGCTGTCTGTCGTGTGTAGGACTGTAGAGCTGTAGAGGAGGAAGGAACGTGAACG  236254
Qy     1353  TTCCAGAATGTGTGTAAATTGAATAAAGCTAGCCGTGTAAATGATCCTGCATAA  1412
Db   236253  CTCCAGAATGTGTGTAAATCGGATAAAGCTAGAAGTGACCTCAGTCGTGTCGCTGCATAA  236194
Qy     1413  TCTCTTCATTCCGAGGAGCACCCCACCCCCATTCTTAAATTGT  1472
Db   236193  TCTCTT------CGCCCCGGCCCACCCCGTTCCTGAATTGT  236160
```

```
                                              /gene="CXCR-4"
                                              /number=1
                                              2308..3366
                              exon            /gene="CXCR-4"
                                              /number=2

ORIGIN

Query Match          62.2%;  Score 1054.2;  DB 10;  Length 3366;
Best Local Similarity 99.7%;  Pred. No. 7.8e-221;
Matches 1056;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy    1 ATATACACTTCTGATAACTACTCTGAGGAAGTGGGGTCTGGAGACTATGACTCCAACAAG   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2308 ATATACACTTCTGATAACTACTCTGAGGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 2367

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCATTCAATAGATCTTCCTGCCACCATCTAC      120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2368 GAACCCTGCTTCCGGGATGAAAACGTCATTCAATAGATCTTCCTGCCACCATCTAC     2427

Qy  121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGGTTAC    180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2428 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGGTTAC   2487

Qy  181 CAGAAGAGCTAAGGACATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC    240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2488 CAGAAGAGCTAAGGACATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC   2547

Qy  241 CTCTTGTCATCACACTCCCTTCTGGGCAGTGATGCATGGCAGTTGACTGGTACTTTGGG   300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2548 CTCTTTGTCATCACACTCCCTTCTGGGCAGTTGATGCATGGCAGTTGACTGGTACTTTGGG 2607

Qy  301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACGTCAACCTCTACAGACAGCGTTCTC   360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2608 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTACACGTCAACCTCTACAGACAGCGTTCTC 2667

Qy  361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCCGCCATTGTCCACGCCACCAACAGTCAA  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2668 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCCGCCATTGTCCACGCCACCAACAGTCAA 2727

Qy  421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2728 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 2787

Qy  481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGGGGACATCAGTCAGGGG    540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2788 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGGGGACATCAGTCAGGGG   2847

Qy  541 GATGACAGTACATCTGACGGGTCTCATCCTCCCGATAGCCTGTGATGGTGGTTTCAA    600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2848 GATGACAGTACATCTGACGGGTCTCATCCTCCCGATAGCCTGTGATGGTGGTTTCAA   2907

Qy  601 TTCCAGCATATAATGTGGGTCTCATCCTCGTCACTCGCCTGCTGCTGCCATATTATGTGGGGATCAGC  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2908 TTCCAGCATATAATGTGGGTCTCATCCTCGTCACTCGCCTGCTGCTGCCATATTATGTGGGGATCAGC 2967

Qy  661 ATCATCATCCTAAGCTGTCACACTCCAAGGCCTCCAAGGCCCACCAGAAGCCACCAGAGACG  720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2968 ATCATCATCCTCAAGCTGTCACACTCCAAGGCCTCCAAGGCCCACCAGAAGCCACCAGACG 3027

Qy  721 ACAGTGGATCTCCATCCTAGCCTTCTTTGCTGCTGCCTGCATATTATGTGGGGATCAGC   780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3028 ACAGTGGATCTCCATCCTAGCCTTCTTTGCTGCTGCCTGCATATTATGTGGGGATCAGC  3087

Qy  781 ATCGACTCCTTCATCCTTTTGGGAGTGACTTCAGCAAGAGATGTGACTTCGAGAGCATTGTG   840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3088 ATCGACTCCTTCATCCTTTTGGGGTCATCAAGCAAGGATGTGACTTCGAGAGCATCGTG   3147

Qy  841 CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC   900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3148 CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC  3207

Qy  901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCACATGCCAGTCAACTCCATG   960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy 1473 TTGGTTATGCGTGTGTGATGGTTGTTGGTTTTTTTTGTTGTTGTTGTTGTTGTTTT      1532
        ||||||||||||||||
Db 236159 TTGGTTGTGTGCTGTGTG----------------------GTTTC               236139

Qy 1533 TCTGTAAAGATGGCACTTAAAACCAAAGCTGAAATGGTGGTAGAAATGCTGGGGTTTT   1592
        |||||||||||||||||||||||||||||||||
Db 236138 TCTGTAGAAGATGGCACTTAAAACCAAAGCCTGAAATGGT-ATAGAAATGCTGG------ 236086

Qy 1593 TTTTGTTTGTTTGTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTA  1652
        --------GTTTTTCTTTCCAGTTTTCAAGGGTGGATTGATTTCAGTCCCTACAAATGTA
Db 236085 --------GTTTTTCTTTCCAGTTTTCAAGGGTGGATTGATTTCAGTCCCTACAAATGTA 236034

Qy 1653 CAGTTCTTGTATTACATTGTTAATAAAAGTCAATGATAAACTT                 1694
        ||||||||||||||||||||||||||||||||||||||||||
Db 236033 CAGTTCTTGTATTATGTGTTAATAAAAGTCAATGGTAAACTT                  235992

RESULT 13
MMU65580
LOCUS       MMU65580                3366 bp    DNA     linear   ROD 13-DEC-1996
DEFINITION  Mus musculus fusin (CXCR-4) gene, complete cds.
ACCESSION   U65580
VERSION     U65580.1  GI:1731650
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3366)
  AUTHORS   Heesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
  TITLE     Cloning of the mouse fusin gene, homologue to a human HIV-1
            co-factor
  JOURNAL   J. Immunol. 157 (12), 5455-5460 (1996)
  MEDLINE   97113334
   PUBMED   8955194
REFERENCE   2  (bases 1 to 3366)
  AUTHORS   Heesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-1996) Pathology, Harvard Medical School, 200
            Longwood Avenue, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..3366
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
     gene            1..3366
                     /gene="CXCR-4"
     CDS             join(1..21,2308..3366)
                     /gene="CXCR-4"
                     /note="SDF-1 chemokine receptor"
                     /codon_start=1
                     /product="fusin"
                     /protein_id="AAC52953.1"
                     /db_xref="GI:3645915"
                     /translation="MEPISVSIYTSDNYSEVGSGDYDSNKEPCFRDENVHFNRIFLP
                     TIYFIIFLTGIVGNGLVLVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMA
                     DWYFGKFLCKAVHIIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKAVYV
                     GWIPALLLTIPDFIFADVSGDISQGDDRYICDRLYPDSLWMVFQFQHIMVGLILP
                     GIVILSCYCIIISKLSHSKGHQKRKALKTTVILIIAFFACWLPYYVGISIDSFILLGV
                     IKQGCDFESIVHKWISITEALAFFHCCLNPILYAFLGAKFKSSAQHALNSMSRGSSLK
                     ILSKGKRGGHSSVSTSESSSFHSS"
     exon            1..21
                     /gene="CXCR-4"
                     /number=1
     exon            1..15
                     /gene="CXCR-4"
                     /note="alternate exon 1"
     intron          16..>2307
                     /gene="CXCR-4"
                     /note="alternate intron 1"
     intron          22..2307
                     /gene="CXCR-4"
                     /note="alternate intron 1"
```

RESULT 14

RNU54791

LOCUS   RNU54791   1362 bp   mRNA   linear   ROD 14-SEP-2001
DEFINITION Rattus norvegicus chemokine receptor LCR1 mRNA, complete cds.
ACCESSION   U54791
VERSION   U54791.1 GI:1354504
KEYWORDS   .
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 1362)
  AUTHORS   Wong,M.L., Xin,W.W. and Duman,R.S.
  TITLE   Rat LCR1: cloning and cellular distribution of a putative chemokine
            receptor in brain
  JOURNAL   Mol. Psychiatry 1 (2), 133-140 (1996)
  MEDLINE   97233286
   PUBMED   9118323
REFERENCE   2 (bases 1 to 1362)
  AUTHORS   Wong,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (11-APR-1996) Ma-Li Wong, Clincial Neuroendocrinology
            Branch, National Institutes of Health, NIMH, Building 10 Room
            3S231, 10 Center DR MSC 1284, Bethesda, MD 20892-1284, USA
FEATURES   Location/Qualifiers
     source   1..1362
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="RLCR1"
     CDS   44..1042
                     /codon_start=1
                     /product="chemokine receptor LCR1"
                     /protein_id="AAB01981.1"
                     /db_xref="GI:1354505"
                     /translation="MEIYTSDNYSEEVGSGDYDSNKEPCFRDENENFNRIPLPTIYFI
                     IFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFSDAMADWYFGKF
                     LCKAVHIIYTVNLYSSVLILAFISLDRYLA1VHATNSQSARKLLAEKADIFADVSQG
                     DGRYICDRLYPDSLMWVFQRQHIMVGLILPGIVILSCYCIIISKLSHSKGHQKRKAL
                     KTIVILILAFFACWLPYYVGISIDSFILLEVIKQGCEFESVVHKWISITEALAFFHCC
                     LNPILYAFLGAKFKTSAQHALNSMSRGSSLKILSKGKRGGHSSVSTESESSSFHSS"

ORIGIN

Query Match   55.9%; Score 946.4; DB 10; Length 1362;
Best Local Similarity   84.8%; Pred. No. 3.6e-197;
Matches 1169; Conservative   0; Mismatches 116; Indels 93; Gaps   6;

```
Db   1244 TTTTTCTGTTGTTGTCTGTTCACGTGAATGAGTGTCTAGGCAGGACCTGTGGCCAAGT 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1294 TCTTAGTAGCTGTTATCTGTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAAC 1351
          ||||| || || |||||
Db   1304 TCTTAGTCGCTGTGTGTCTGTGTGTAGGACTGTAGAGCTGTAGAGGAAGGAACTGAAC 1361

RESULT 15
AX481810
LOCUS        AX481810        1582 bp    DNA        linear    PAT 16-AUG-2002
DEFINITION   Sequence 11 from Patent WO02057308.
ACCESSION    AX481810
VERSION      AX481810.1    GI:22316633
KEYWORDS
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
  AUTHORS    Demaison,C., Isalan,M., Moore,M., Choo,Y., Girdlestone,J.R. and
             Ullman,C.
  TITLE      Nucleic acid binding polypeptides
  JOURNAL    Patent: WO 02057308-A 11 25-JUL-2002;
             SANGAMO BIOSCIENCES INC (US)
FEATURES             Location/Qualifiers
     source          1..1582
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match      55.1%;    Score 933.4;    DB 6;    Length 1582;
Best Local Similarity   81.8%;    Pred. No. 2.6e-194;
Matches 1154;    Conservative    0;    Mismatches 216;    Indels    41;    Gaps    5;

Qy      1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
          |||||||||| | ||||||| |||||||| || |||||||||||||| ||||||||||||
Db      7 ATATACACTTCAGATAACTACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCATGAAG 66

Qy     61 GAACCCTGCTTCCGGGATGAAAAGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 120
          ||||| || |||||| ||||||||||||| ||||||||| ||| || ||||||||||||
Db     67 GAACCCTGTTTCCGTGAAGAAAATGTAATTCAATAAAAATCTTCCTGCCACCATCTAC 126

Qy    121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
          || |||||||| ||| |||| | || || ||||| |||| ||||||| || || ||||
Db    127 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 186

Qy    181 CAGAAGAAGCTAAGGAGCAATGACGGACAAGTACCGCTGCACCTTCAGTGGCTGACCTC 240
          | ||||||||||||||||| ||| ||||| ||||||||||| || || |||||| ||||
Db    187 CAGAAGAAACTGAGAGAACATGACGGACAAGTACCAGGCTGCACCTGCACCTGGTTGGACCTC 246

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGCAGTTGATGCCATACACTGTCAACCTCTACAGCAGCGGGTTCTC 300
          |||||||||||||| |||||||| ||||||||||||||| |||||| || ||||||| ||||||||||| |||
Db    247 CTCTTTGTCATCATCACGCCTCCTCTGGCAGTTCATCACACAGTCAACCTCTACAGCAGTGTCCCTC 306

Qy    301 AAATTTTGTGTGAAGGCTGTCCATATCATCATCACACTGTCCACGCCACCACCAAGTGCA 360
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    307 AACTTCCATGCAAGGCAGTCCATGTCATCATCACACAGCCAGTGTCCTC 366

Qy    361 ATCCTGGCCTTCATCAGCTGGACCCTGACCCGTACCTGCCATTGTCCACGCCACCACCAA 420
          |||||||||||||||||||||||||||||| ||||||||||| ||| ||| ||||||| |
Db    367 ATCCTGGCCTTCATCAGTCTGGACCCTGACCCCGTGGCGTCTGGATCCCAGCCCTC 426

Qy    421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGAATCCCTGCCCTC 480
          ||||||||||||| ||||||||||| |||||| ||||| ||| |||||||||||||||||
Db    427 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTTCATGTTGCGTCTCTGGATCCCTGCCCTC 486

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGGGACATCAGTCAGGGG 540
          ||| |||||| |||||| ||||||| |||||||||||||||||||| ||||||
Db    487 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG--------TCAGTGAGGGCA 531

Qy    541 GATGACAGGTACATCGTGACCCGCCCTTACCCGGATGAGCCTGTGGATGGTGTTTCAA 600
          ||||||| | ||||||||||||||||| | |||||| ||||||
```

Search completed: May 17, 2004, 12:10:55
Job time : 6860.38 secs

```
Db    532 GATGACAGATATATCTGTGACCGCCTTCACCCCAATGACTTGTGGGTGGTTGTGTTCCAG 591
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy    601 TTCCAGCATATAATGGTGGGTCTCATGCGGACTGTCATCCTCTCCTGTTACTGC 660
          |||||||||||| ||||||
Db    592 TTTCAGCACATCATCATGGTGGCCTTATCCTGGTATTGTCATCCTGTCCTGCTATTGC 651

Qy    661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGGCCACCAGAACGCAAGGGCCCTCAAGACG 720
          || ||||||||| |||||||||||||||||| |||||||||||||||||||| ||||| |
Db    652 ATATCATCTCCAAGCTGTCACACTCCAAGGGCCAAGGGCGCAAGGCCCTCAAGACC 711

Qy    721 ACAGTCATCCTCATCCTAGCTTTCTTGCCTGGCTGCCATATTATGGGGGATCAGC 780
          |||||||||||||||||||||||||||||||||||| ||||||||| ||||||||||
Db    712 ACAGTCATCCTCATCCTGGCTTTCTTGCCTGGCTGGGCCTTACTACATTGGGATCAGC 771

Qy    781 ATGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGGAGAGCATTGTG 840
          ||||||||||||||||||| |||||||||||||||| || ||| |||| |||||| |||
Db    772 ATCGACTCCTTCATCCTCTGGAAATCATCAAGCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 831

Qy    841 CACAAGTGGATCCATCACAGAGGCCCTCGCCTTCTCCACTGTTGCTGCTGACCCCATC 900
          |||||||||||||||||||||||||||| |||||| |||||||||||||||||| |||
Db    832 CACAAGTGGATTTCATCACCGAGGCCCTAGCTTTCTTCCACTGTTGCTGCTGACCCCATC 891

Qy    901 CTCTATGCCTTCCTCGGGCCAAGTTCAAAGCTCTGCCCAGCATGACTCAACTCCATG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    892 CTCTATGCTTTCCTTGGAGCCAAATTTAAAACCTGCCCAGCACGCACTCACCTCTGTG 951

Qy    961 AGCAGAGGCTCCAGCCTCAGATCCTTTCCAAAGGAAGCGGGGTGGACACTCTTCCGTC 1020
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    952 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAGCAGGTGGACATTCATCGTGTT 1011

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCCAACGACTTATAT 1080
          |||||||||||||||||||||||||||||||||||                          ||
Db   1012 TCCACTGAGTCTGAGTCTTCAAGTTTTCACTCCAGCTAA-------CACAGATGTAAAA 1063

Qy   1081 AATATATATAATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAGAG 1140
          |||||||||||||||||||||||||||||||| |||| ||||| |||||||||||
Db   1064 GACTTTTTTTATACGATAAAATAACTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1123

Qy   1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTAATTGACTGTTGGAGTTTATGTTCC 1200
          ||||||||||||| |||||| ||||||||||| ||      ||||||||||||
Db   1124 ACTGACCAATATGTACAGTTTTTAT GCTTGTGTTGGATTTTTGT-------CTTGTGTTTC 1177

Qy   1201 TCTAGTTTTTGTGAGGTTGACTTAATTATATAAAATATGTTTTTGTTTGTTTCATGT 1260
          |||||||||||||||||| |||||||  |||||| ||||||||||
Db   1178 TTTAGTTTTTGTGAAGTTAATTGACTATTTATA----TAAATTTTTTGTTCATAT 1233

Qy   1261 GAATGAGCCTCAGGCAGCAGACCTGTCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
          |||||||| ||| || ||| || ||||||||||||||| ||||| ||| ||||||||||
Db   1234 TGATGTGTGTCAGGCAGGCAGACCTGTCATCCTTCTTAGTTGCTCTTCGTGGTAG 1293

Qy   1321 GACTGTAGAACTGTAGAAGAGGAAGAAACTGAACATTCCAGAAATGTGTGTAAAATGAATAAA 1380
          |||||||                                   |||||||||||||||||||||||
Db   1294 GACTGTAGAA--------AAGGAACTGAACATTCCAGAGCAGTGAATCACGTAAA 1345

Qy   1381 GCTAGCCCGTGATCCTCAGCTGTTGCTGCTATA 1411
          ||||||||||||||||||||||||||||||
Db   1346 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1376
```

```
                    GenCore version 5.1.6

          Copyright (c) 1993 - 2004 Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:      May 17, 2004, 03:46:58 ; Search time 726.44 Seconds
                                        (without alignments)
                                     9906.453 Million cell updates/sec

Title:          US-09-367-052-7
Perfect score:  1694
Sequence:       1 atatacacttctgataacta.........taaaagtcaatgataaactt 1694

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:        6747726


Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      N_Geneseq_29jan04:*
                1: geneseqn1980s:*
                2: geneseqn1990s:*
                3: geneseqn2000s:*
                4: geneseqn2001as:*
                5: geneseqn2001bs:*
                6: geneseqn2002s:*
                7: geneseqn2003as:*
                8: geneseqn2003bs:*
                9: geneseqn2003cs:*
               10: geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1694	100.0	1758	2	AAZ27611	Aaz27611 Mouse CXC
2	1692.4	99.9	1877	2	AAV46370	Aav46370 Nucleic a
3	933.4	55.1	1582	6	ABS64793	Abs64793 CXCR4 rec
4	933.4	55.1	1588	2	AAZ27610	Aaz27610 Human CXC
5	933.4	55.1	1664	3	AAA61656	Aaa61656 Human CXC
6	933.4	55.1	1664	3	AAZ40014	Aaz40014 CXCR4 cod
7	933.4	55.1	1679	4	AAD12801	Aad12801 Human neu
8	933.4	55.1	1679	4	AAI65467	Aai65467 Nucleotid
9	933.4	55.1	1679	6	ABS53992	Abs53992 DNA encod
10	933.4	55.1	1679	6	ABV78039	Abv78039 Hypoxia-r
11	933.4	55.1	1679	7	ABX08779	Abx08779 Angiogene
12	933.4	55.1	1679	7	ABX74454	Abx74454 Human cDN
13	933.4	55.1	1679	7	ABZ68886	Abz68886 Nucleotid
14	933.4	55.1	1679	7	ABZ42642	Abz42642 Human CXC
15	933.4	55.1	1679	9	ADC98645	Adc98645 Human CXC
16	933.4	55.1	1711	6	ABZ35348	Abz35348 Human gen
17	933.4	55.1	1711	9	ADB47320	Adb47320 Human cDN
18	933.4	55.1	1737	2	AAQ80521	Aaq80521 Human mon
19	933.4	55.1	1737	2	AAQ99007	Aaq99007 Chemokine
20	933.4	55.1	1944	2	AAX15882	Aax15882 cDNA enco
21	933.4	55.1	5161	3	AAA34774	Aaa34774 Human ade
22	933.4	55.1	5161	6	AAF20896	Aaf20896 Human CCR
23	933.4	55.1	5161	6	AAH99951	Aah99951 CXCR4 gen
24	933.4	55.1	5161	7	ABZ96590	Abz96590 Human CCR
25	933.4	55.1	5161	7	ACA64841	Aca64841 Human CXC
26	931.8	55.0	8747	9	ADD67548	Add67548 Human Ly1
27	931.4	55.0	1670	6	ABK83803	Abk83803 Human cDN
28	931.4	55.0	1670	6	ABN95645	Abn95645 Gene #214
29	931.4	55.0	1670	7	ACC46765	Acc46765 Human COP
30	931.4	55.0	1679	6	ABS53750	Abs53750 DNA_encod
31	930.2	54.9	1737	9	AAQ29506	Aaq29506 New plate
32	905.4	53.4	1050	9	ADB58421	Adb58421 Toxicity-
33	905.4	53.4	1050	9	ADB53005	Adb53005 Primary r
34	855.6	50.5	1225	9	ABZ35630	Abz35630 Human gen
35	855.6	50.5	1225	7	ACA56637	Aca56637 Human sig
36	843.6	49.8	1224	6	ABZ35563	Abz35563 Human gen
37	834	49.2	1317	6	AAV18357	Aav18357 Human RM3
38	834	49.2	1317	3	AAA91726	Aaa91726 Human 7TM
39	834	49.2	1317	6	ABK54255	Abk54255 Human 7 t
40	832.4	49.1	1317	6	AAQ66179	Aaq66179 Seven tra
41	806.2	47.6	1059	3	AAZ38553	Aaz38553 Human CXC
42	806.2	47.6	1059	6	ABQ74938	Abq74938 Human CXC
43	806.2	47.6	1059	6	AAH99952	Aah99952 CXCR4 enc
44	804.6	47.5	1059	7	ACC72707	Acc72707 Human can
45	804.6	47.5	1102	7	ACA64750	Aca64750 Chemokine

ALIGNMENTS

```
RESULT 1
AAZ27611
ID   AAZ27611 standard; DNA; 1758 BP.
XX
AC   AAZ27611;
XX
DT   16-DEC-1999  (first entry)
XX
DE   Mouse CXCR4 coding sequence.
XX
KW   CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;
KW   tissue repairing agent; vascularisation; ss.
XX
OS   Mus sp.
XX
PN   WO9948528-A1.
XX
PD   30-SEP-1999.
XX
PF   23-MAR-1999;     99WO-JP001448.
XX
PR   24-MAR-1998;     98JP-00095448.
XX
PA   (CHUS ) CHUGAI SEIYAKU KK.
PA   (KISH/) KISHIMOTO T.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K;
XX
DR   WPI; 1999-591042/50.
DR   P-PSDB; AAY39994.
XX
PT   CXCR4-potentiating agents and methods useful for inhibiting
PT   neovascularization, and treating solid cancers.
XX
PS   Disclosure; Page 50-51; 63pp; Japanese.
XX
CC   This sequence encodes the mouse CXCR4 protein. The invention relates to
CC   remedies inhibiting neovascularisation, remedies for solid cancer,
CC   remedies for diseases pathologically caused by neovascularisation and
CC   tissue repairing agents containing as the active ingredient a substance
CC   capable of potentiating CXCR4. Based on a finding that vascularisation is
CC   inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC   remedies inhibiting vascularisation which contain as the active
CC   ingredient a substance capable of potentiating CXCR4, remedies for solid
CC   cancer, remedies for diseases pathologically caused by neovascularisation
CC   and tissue repairing agents containing as the active ingredient a
```

CC   substance capable of potentiating CXCR4. It is also possible to establish
CC   methods for treatment with the use of these remedies
XX
SQ   Sequence 1758 BP; 447 A; 414 C; 369 G; 528 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1694;   DB 2;   Length 1758;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1694;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAACGTGGGGTCTGAGACTATGACTCCAACAAG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   22 ATATACACTTCTGATAACTACTCTGAAGAACGTGGGGTCTGAGACTATGACTCCAACAAG   81

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCATTTCAATAGGATCTTCCTGCCACCATCTAC   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   82 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC   141

Qy   121 TTCATCATCTTCTTGACTGGCATAGTGGCATGATTGGTGATCCTGGTCATGGGTTAC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   142 TTCATCATCTTCTTGACTGGCATAGTGGCATGATTGGTGATCCTGGTCATGGGTTAC   201

Qy   181 CAGAAGAAGCTAAGGAGCATGACGGACAAGCTACCGGCTGCACCTGTCAGTGGCTGACCTC   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   202 CAGAAGAAGCTAAGGAGCATGACGGACAAGCTACCGGCTGCACCTGTCAGTGGCTGACCTC   261

Qy   241 CTCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGATGCCATGGCTGACTGTTACTTTGGG   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   262 CTCTTTGTCATCACACTCCCCTTCTCGGGCAGTTGATGCCATGGCTGACTGTTACTTTGGG   321

Qy   301 AAATTTTTGTGTAAGGCTGTCCATATCATCTCAACCTCTACACAGCAGCCTTCTC   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 AAATTTTTGTGTAAGGCTGTCCATATCATCTCAACCTCTACACAGCAGCCTTCTC   381

Qy   361 ATCCTGGCCTTCATCAGCCTGACCGGTACCGGCGCCATTGTCCACCGCCACCAACAGTCAA   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   382 ATCCTGGCCTTCATCAGCCTGACCGGTACCGGCGCCATTGTCCACCGCCACCAACAGTCAA   441

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   442 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC   501

Qy   481 CTCCTGACTATACCTGACACTTCATCTTTGCCACGCTCAGCAGCCAGGGGACACATCAGTCAGGGG   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   502 CTCCTGACTATACCTGACACTTCATCTTTGCCACGCTCAGCAGCCAGGGGACACATCAGTCAGGGG   561

Qy   541 GATGACAGGTACATCGTGACCGGCCTTTACCCGGATAGCCTGTGGATGGTGGTGTTTCAA   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   562 GATGACAGGTACATCGTGACCGGCCTTTACCCGGATAGCCTGTGGATGGTGGTGTTTCAA   621

Qy   601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCCGGCGCATCGTCATCCTCTCTGTTACTGC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   622 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCCGGCGCATCGTCATCCTCTCTGTTACTGC   681

Qy   661 ATCATCATCTCTAAGCTGTCACACTGTCACACTCCAGAGCGCCACCAGAGACGCGCAAGGCCTCAAGGACACG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   682 ATCATCATCTCTAAGCTGTCACACTGTCACACTCCAGAGCGCCACCAGAGACGCGCAAGGCCTCAAGGACACG   741

Qy   721 ACAGTGGATCTCCATCCCAGCTTCTTTGCCTGGCTGGCTGCCATATTATGTGGGGATCAGC   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   742 ACAGTGGATCTCCATCCCAGCTTCTTTGCCTGGCTGGCTGCCATATTATGTGGGGATCAGC   801

Qy   781 ATCGACTCCTTCATCCTTTGGGAGTCATCATCAAGCAAGGATGGACTTGACTTCGAGAGCATTGTG   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   802 ATCGACTCCTTCATCCTTTGGGAGTCATCATCAAGCAAGGATGGACTTGACTTCGAGAGCATTGTG   861

Qy   841 CACAAGTGGATCCATCAGAGGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCCATC   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   862 CACAAGTGGATCCATCAGAGGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCCATC   921

Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   922 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG   981

Qy   961 AGCAGAGGCTCCAGCCTCCAAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   982 AGCAGAGGCTCCAGCCTCCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC   1041

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1042 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT   1101

Qy   1081 AATATATATATATATGATAATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 AATATATATATATATGATAATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG   1161

Qy   1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC   1221

Qy   1201 TCTAGTTTTGTGAGGTTGACTTAATTATATAAAATATTGTTTTTGTTTTTGTTTCATGT   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 TCTAGTTTTGTGAGGTTGACTTAATTATATAAAATATTGTTTTTGTTTTTGTTTCATGT   1281

Qy   1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG   1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG   1341

Qy   1321 GACTGTAGAACTGTAGAGGAGAGAAACTGAACATTCCAGAAGTGTGTGTAAATTGAATAAA   1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1342 GACTGTAGAACTGTAGAGGAGAAGAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAA   1401

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCCATAATCTCTCCATTCCGAGGAGCACCCACCC   1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1402 GCTAGCCGTGATCCTCAGCTGTTGCTGCCATAATCTCTCCATTCCGAGGAGCACCCACCC   1461

Qy   1441 CCACCCCCACCCCACCCCATTCTTAAATTGTTTGGTATGCTGTGTGATGGTTTGTTTG   1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1462 CCACCCCCACCCCACCCCATTCTTAAATTGTTTGGTATGCTGTGTGATGGTTTGTTTG   1521

Qy   1501 GTTTTTTTGTTGTTGTTGTTGTTTTTTTTTTTTCTGTAAAAGATGGCACTTAAAACCAAA   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1522 GTTTTTTTTGTTGTTGTTGTTGTTTTTTTTTTTCTGTAAAAGATGGCACTTAAAACCAAA   1581

Qy   1561 GCCTGAAATGGTGGTAGAAATGCTGGGGGTTTTTTGTTTTGTTTTTTTTCAGTTTTCA   1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1582 GCCTGAAATGGTGGTAGAAATGCTGGGGGTTTTTTGTTTTGTTTTTTTTCAGTTTTCA   1641

Qy   1621 AGAGTAGATTGACTTCAGTCCCTACAAAATGTACAGTCTTCGTGTATTACAATTGTTAATAAAAG   1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1642 AGAGTAGATTGACTTCAGTCCCTACAAAATGTACAGTCTTCGTGTATTACAATTGTTAATAAAAG   1701

Qy   1681 TCAATGATAAACTT   1694
          ||||||||||||||
Db   1702 TCAATGATAAACTT   1715

RESULT 2
AAV46370
ID   AAV46370 standard; cDNA to mRNA; 1877 BP.
XX
AC   AAV46370;
XX
DT   20-NOV-1998 (first entry)
XX
DE   Nucleic acid encoding a murine CXC chemokine receptor.
XX
KW   Mouse; CXC chemokine receptor; pre-B cell line DW34;
KW   CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1; HIV infection;
KW   screening; inhibitor; AIDS; ds.
XX
OS   Mus sp.
XX
FH   Key   Location/Qualifiers
FT   CDS   120..1199
FT   /*tag= a
XX

PN   WO9835035-A1.
XX
PD   13-AUG-1998.
XX
PF   07-FEB-1997;   97WO-JP000299.
XX
PR   07-FEB-1997;   97WO-JP000299.
XX
PA   (SHIO) SHIONOGI & CO LTD.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;
PI   Nakajima T, Yoshie O;
XX
DR   WPI; 1998-447232/38.
DR   P-PSDB; AAW64778.
XX
PT   Mouse CXC chemokine receptor binding to PBSF/SDF-1 pre-B cell stimulatory
PT   factor - is useful for screening of potential HIV infection and AIDS
PT   inhibitors.
XX
PS   Claim 3; Page 39-42; 76pp; Japanese.
XX
CC   The present sequence encodes a murine CXC chemokine receptor which binds
CC   to the mouse CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1. The
CC   nucleic acid is isolated from mouse pre-B cell line DW34. The receptor
CC   and cells expressing it can be used in the study and mapping of the
CC   mechanism of HIV infection and in screening of potential inhibitors of
CC   HIV infection and the development of AIDS
XX
SQ   Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T; 0 U; 0 Other;

Query Match   99.9%; Score 1692.4; DB 2; Length 1877;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 1693; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy      1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 200

Qy     61 GAACCCTGTTCCCGGGATGAAAACGTCCATTCAATAGGAAGTCCTTCCTGCCCACCATCTAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 GAACCCTGTTCCCGGGATGAAAACGTCCATTCAATAGGAAGTCCTTCCTGCCCACCATCTAC 260

Qy    121 TTCATCATCTTCTTGACTAGTGGCATAGTGGCATAGTCGGCAATGGATTGGTGATCCTGGT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 TTCATCATCTTCTTGACTAGTGGCATAGTGGCATAGTCGGCAATGGATTGGTGATCCTGGT 320

Qy    181 CAGAAGAAGCTAAGGAGGAGCATGACGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 CAGAAGAAGCTAAGGAGGAGCATGACGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTG 380

Qy    241 CTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCCATGGCTGGTACTTTGGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 CTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCCATGGCTGGTACTTTGGG 440

Qy    301 AAATTTTGTGTAAGGCTGTCATATCATCACTGTCAACCTCACAGCAGCAGCGTCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 AAATTTTGTGTAAGGCTGTCATATCATCACTGTCAACCTCACAGCAGCAGCGTCTC 500

Qy    361 ATCCTGGCTTCATCAGCTGGACCGGTACCTGCACGGCCACCACCAGTCAA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 ATCCTGGCTTCATCAGCTGGACCGGTACCTGCACGGCCACCACCAGTCAA 560

Qy    421 AGGCCAAGGAAACTGCTGGCTGGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    561 AGGCCAAGGAAACTGCTGGCTGGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 620

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCAGCCAGGGGACATCAGTCAGGGGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    621 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCAGCCAGGGGACATCAGTCAGGGGG 680

Qy    541 GATGACAGGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGATGGTGTGTTCAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    681 GATGACAGGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGATGGTGTGTTCAA 740

Qy    601 TTCCAGCATATAATGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTGTTACTGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    741 TTCCAGCATATAATGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTGTTACTGC 800

Qy    661 ATCATCATCCTCTAAGCTGTCACACTCCAAGGGCCACCAGAGCGCCTCAAGAGACG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    801 ATCATCATCCTCTAAGCTGTCACACTCCAAGGGCCACCAGAGCGCCTCAAGAGACG 860

Qy    721 ACAGTGCATCCTCATCCTAGCTTCTTTGCCTGCTGCTCCATATTATGTGGGATCAGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    861 ACAGTCATCCTCATCCTAGCTTCTTTGCCTGCTGCTCCATATTATGTGGGGATCAGC 920

Qy    781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 900

Qy    841 CACAAGTGGATCTCCATCACAGAGGGCCCTGCCTTCTTCCCACTGTTGCCTGAACCCCATC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    981 CACAAGTGGATCTCCATCACAGAGGGCCCTGCCTTCTTCCCACTGTTGCCTGAACCCCATC 1040

Qy    901 CTCTATGCCTTCCTGCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1041 CTCTATGCCTTCCTGCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG 1100

Qy    961 AGCAGAGGCTCCAGCCTCAAGATCCTTCAAAGGAAAGGCGGGGTGGACACTCTTCCGTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1101 AGCAGAGGCTCCAGCCTCAAGATCCTTCAAAGGAAAGGCGGGGTGGACACTCTTCCGTC 1160

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTAACCCTTATGCAAAGACTTATAT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1161 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTAACCCTTATGCAAAGACTTATAT 1220

Qy   1081 AATATATATATGATAAAGAACTTTTTATTGTTACACATTTTCCAGATATAAGAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221 AATATATATATGATAAAGAACTTTTTATTGTTACACATTTTCCAGATATAAGAG 1280

Qy   1141 ACTGACCAGTCTGTACAGTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1281 ACTGACCAGTCTGTACAGTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1340

Qy   1201 TCTAGTTTTGTGAGGTTGACTTAATTTATTATAAAATTGTTTTTGTTTGTTTCATGT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1341 TCTAGTTTTGTGAGGTTGACTTAATTTATTATAAAATATTGTTTTTGTTTGTTTCATGT 1400

Qy   1261 GAATGAGCGTCTAGGCAGGACCTGTGCCAAGTTCTTAGTAGCTGTTATCTGTGTGTAG 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1401 GAATGAGCGTCTAGGCAGGACCTGTGCCAAGTTCTTAGTAGCTGTTATCTGTGTGTAG 1460

Qy   1321 GACTGTAGAACTGTGAGGAGAAACTGAACATTCAGAATGTGTGGTAAATTGAATAAA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1461 GACTGTAGAACTGTGAGGAGAAACTGAACATTCAGAATGTGTGGTAAATTGAATAAA 1520

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGTGCTGCATAATCTTCTTCATTCCGAGGAGCACCCCACCC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1521 GCTAGCCGTGATCCTCAGCTGTTGTGCTGCATAATCTTCTTCATTCCGAGGAGCACCCCACCC 1580

Qy   1441 CCACCCCACCCCATTCTTAAATTGTTGGTTATGCTGTGTGATGGTTGTTTG 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1581 CCACCCCACCCCATTCTTAAATTGTTGGTTATGCTGTGTGATGGTTGTTTTG 1640

Qy   1501 GTTTTTTTTGTTGTGTGTGTTGTTGTTGTTTTCTGTAAAAGATGGCACTTAAAACCAA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1641 GTTTTTTTTGTTGTGTGTGTTGTTGTTGTTTTCTGTAAAAGATGGCACTTAAAACCAA 1700

Qy   1561 GCCTGAAATGGTGGTAGAAATGCTGGGGTTTTTTTTGTTTTTGTTTTTTCAGTTTTCA 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1701 GCCTGAAATGGTGGTAGAAATGCTGGGGTTTTTTTGTTTTTGTTTTTTCAGTTTTCA 1760

Qy   1621 AGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTGATTACATTGTTAATAAAAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1761 AGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTGATTACATTGTTAATAAAAG 1820
```

Qy   1681 TCAATGATAAACTT 1694
           |||||||||||||
Db   1821 TCAATGATAAACTT 1834

RESULT 3
ABS64793
ID   ABS64793 standard; DNA; 1582 BP.
XX
AC   ABS64793;
XX
DT   15-NOV-2002 (first entry)
XX
DE   CXCR4 receptor DNA sequence.
XX
KW   Receptor; virus infection; immune response; autoimmune disease;
KW   TNFR1 receptor; inflammation; autoimmune encephalomyelitis;
KW   rheumatoid arthritis; myocarditis; CXCR4 receptor; ds.
XX
OS   Unidentified.
XX
PN   WO20025730B-A2.
XX
PD   25-JUL-2002.
XX
PF   22-JAN-2002; 2002WO-GB000246.
XX
PR   22-JAN-2001; 2001GB-00001576.
PR   07-FEB-2001; 2001GB-00003032.
XX
PA   (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI   Moore M, Isalan M, Reynolds L, Ullman C, Girdlestone J;
PI   Demaison C, Choo Y;
XX
DR   WPI; 2002-590720/63.
XX
PT   New polypeptide capable of binding nucleic acids for treating or
PT   preventing a disease caused by a virus or a disease associated with an
PT   immune response, particularly autoimmune diseases e.g. inflammation or
PT   rheumatoid arthritis.
XX
PS   Disclosure; Page 12-13; 105pp; English.
XX
CC   The present invention relates to a new polypeptide capable of binding to
CC   a nucleic acid comprising a receptor nucleotide sequence which can
CC   function as a receptor for virus infection or is involved in an immune
CC   response. The molecules of the invention are useful in the treatment or
CC   prevention of a disease caused by a virus or of a disease associated with
CC   an immune response, particularly an autoimmune disease. The zinc finger
CC   polypeptide or a nucleic acid encoding such a polypeptide can be used to
CC   modulate transcription of a receptor nucleotide sequence. The nucleic
CC   acid polypeptide capable of binding to a nucleic acid sequence comprising
CC   a receptor nucleotide sequence, where the receptor is capable of
CC   functioning as a receptor for infection by the virus or is involved in an
CC   immune response, or a nucleic acid encoding such a polypeptide, can be
CC   used in the preparation of a medicament for use in the treatment or
CC   prevention of a disease caused by a virus or of a disease associated with
CC   an immune response. The polypeptides may further be used to treat or
CC   prevent various diseases or syndromes associated with or caused by
CC   malfunction of the receptor as the TNFR1 receptor, such as autoimmune
CC   diseases including inflammation, autoimmune encephalomyelitis, rheumatoid
CC   arthritis or myocarditis. The present nucleic acid sequence represents a
CC   CXCR4 receptor sequence, as described in the invention
XX
SQ   Sequence 1582 BP; 386 A; 376 C; 333 G; 487 T; 0 U; 0 Other;

Query Match   55.1%; Score 933.4; DB 6; Length 1582;
Best Local Similarity 81.8%; Pred. No. 4.2e-207;
Matches 1154; Conservative   0; Mismatches 216; Indels 41; Gaps 5;

Qy   1 ATATACACTTCTGATAACTACTCTGAGAAGTGAGGGTCTGAGACTATGACTCCAACCAAG 60

           ||||||||||| | || ||||||||||||||||||
Db   7 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGACTATGACTCCATGAAG 66

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
           |||||| |||||||| ||| |||||||| || |||| || |||||| ||||||||
Db   67 GAACCCTGTTTCCGTGAGAAGAAATGCTAATTTCATAAAATCTTCCTGCCACCATCTAC 126

Qy   121 TTCATCATCTTCTGACTGGCCATAGTCGGCAATGCGATTGGTGATCCTGGTCATGGGTTAC 180
           |||||||||||| |||||| |||||| ||||| ||| ||||||| ||||| ||||| |||
Db   127 TCCATCATCTTCTTAACTGGCATTGTGGGGCAATGGATTGGTGTCATCCTGGTCATGGGTTAC 186

Qy   181 CAGAAGAAGCTAAGGAGCATGAGCGACATGACCGGCTGCACCGTCAGTGGCTGCGACCTC 240
           ||||||||||| |||||||||| |||||| || |||||||||||||||||||| |||||||
Db   187 CAGAAGAAACTGAGAAGCATGAGCGACATGACAGGCTGCACCGTCAGTGGCCGACCTC 246

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCTGACTGGTACTTGGG 300
           |||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||
Db   247 CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG 306

Qy   301 AAATTTTGTGTAGGCTGTCCATATCATCCACTGTCAACCTCTACAGCAGCGTTCTC 360
           |||| | ||||||||| ||||||||||||| |||||||||||||||||||||||
Db   307 AACTTCCTATGCAAGGCAGTCCATGTCATCACAGTCATCTACAGCAGTGTCCTC 366

Qy   361 ATCCTGGCCCTCATCAGCCTGGACCGGTACCTGCCATTGTCCACGCCACCAACAGTCAA 420
           |||||| ||||||||||| ||||| |||||||||||||||| |||||| ||||||||| |
Db   367 ATCCTGGCCTTCATCAGTCTGGACCGGTACCTGGCCATCGTCACGCCACCAACAGTCAG 426

Qy   421 AGGCCAAGAAAACTGCTGGCAGCAGTCAGTGGGCGTCTGAATCCGATCCGAGCCCCTC 480
           ||||||||| | |||||||||| |||||||||||| |||| |||||||| ||| |||||
Db   427 AGGCCAAGGACGTGTTGGCCGTCAGTGGGTCATGTTGGCGTCTGGATCCCTGCCCTGC 486

Qy   481 CTCCTGACTATACCTGACTTCATCTTGCCGACGTCAGCCGGGGGGACTCAGTCAGTCAGGAG 540
           ||||| ||||||| |||||| ||| | | ||||||||
Db   487 CTGCTGACTATTCCCGACTTCATCTTTGCCAAG----------TCAGTGAGGCA 531

Qy   541 GATGACAGGTACATCTGTGACCGCTTCACCCCGATAGCGTGTGGATGGTGGTGTTTCAA 600
           |||||||| | ||||||||| ||||| ||| ||||||| | ||| ||||||| ||||||
Db   532 GATGACAGATATCGTGACCGCTTCACCCCAATGACTTGTGGGTGGGTTGTTGTGTTCCAG 591

Qy   601 TTCAGCCATATAATGGTGGGTCCATCCGCGCATCGTCATCCTCCTGTTACTGC 660
           |||||||||| |||||||| |||| |||| |||| |||||||||||| ||||||
Db   592 TTTCAGCACATCATGGTTGCCTTATCCTGTGTATTGTCATCCTGCTATTGC 651

Qy   661 ATCATCCTCTAAGCTGTCACACTCCAAGGGCCACCAGAGAGGCCAAGGCCCTCAAGACG 720
           || |||||||| |||||| || |||||||||||||||||||||| |||||||||||
Db   652 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCCAATTATGGGGATCAGC 711

Qy   721 ACAGTCATCCCATCCTAGCTTCTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCCATTGTG 780
           ||||| ||| ||| ||||| |||| || || | |||||| ||| ||| || ||| ||||||
Db   712 ACAGTCATCCTCATCCTGGCTTTCTTCGCCTGTTGGCTTCCTTTACTACATTGGGGATCAG 771

Qy   781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCAATTGTG 840
           || |||||||| |||||| ||| ||||| |||||| |||||||||||||
Db   772 ATCGACTCCTTCATCCTCCTCTGGAAATCATCAAGCAAGGTGTGAGTTTGAGAACACTGTG 831

Qy   841 CACAAGTGGATCTCCATCACAGAGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCATC 900
           ||||| ||||| |||| |||||| |||||| || || || ||||| |||||||||||||
Db   832 CACAAGTGGATTCCATCACCGAGGCCCTCGCCTGCTTTCTTCTTGTCTGAACCCATC 891

Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTCTGCCAAAGTCTGACACTCAACTCCATG 960
           |||| || ||||||| |||||||| ||||||| |||||||||| ||| ||||||
Db   892 CTCTATGCTTTCCTTGGAGCCAAATTTAAACCCTGCCCAAGGAAAGCGGGTGGGTTGGGACTCACCTCTGTG 951

Qy   961 AGCAGAGGCCCAGCCTCAAGATCCTCCAAAGATCCTCCCAAAGATACCCTATCATCTGTT 1020
           |||||| |||| ||| || ||| |||| ||||| |||| |||| | ||| ||| || ||||
Db   952 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGATGAAGCGAGGTGGACACTGTT 1011

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTTTCACTCCAGTTATCCAGCTAAAGACTTATAT 1080
           |||||| |||||||| | | |||| || | |||| |||| | ||| || |
Db   1012 TCCACTGAGTCGAGTCTTCAAGTTTCACTCCAGTCAA------CACAGATGTAAAA 1063

Qy   1081 AATATATATATATATATGATGAAGAAGAACTTTTTATGTTACACATTTCCAGATATAAAGAG 1140
            | |||||||||||| || ||
Db   Mismatches   0;

Query Match   55.1%;   Score 933.4;   DB 2;   Length 1588;
Best Local Similarity   81.8%;   Pred. No. 4.2e-207;
Matches 1154;   Conservative   0;   Mismatches 216;   Indels   41;   Gaps   5;

```
Qy      1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGGTCTGAGACATGAGCCATGACTCCAACAAG 60
          ||||||||||||||||| |||||||||||||||||  ||||||||  ||||||
Db     16 ATATACACTTCAGATAACTACACCGAGGAAATGGGGTCAGGGGACATGAGACTCCATGAAG 75

Qy     61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 120
          ||||||||||||||| || |||||||||||||||||  || |||  ||||||||||||||
Db     76 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCCACCATCTAC 135

Qy    121 TTCATCATCTTCTGACTGGCCATAGTCGGCAATGATTGGTGATCATCGGTTCATGGGTTAC 180
          ||||||||||||||  ||||  |||||||||||||||| |||||||||||||||||||||
Db    136 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGATTGGTCATCCTGGTCATGGGTTAC 195

Qy    181 CAGAAGAAGCTAAGGAGCAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
          ||||||||| |||| || |||||||||||||||||||||||||||| |||||||||||||
Db    196 CAGAAGAAACTGAGAAGCAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC 255

Qy    241 CTCTTTGTCATCACACTTCGGCCCTTCTGGGCAGTTGATGCCATGGCAAACTGGTACTTTGGG 300
          ||||||||||||||||  ||||||| ||||| ||||||||||||||||||||||  ||||||||||
Db    256 CTCTTTGTCATCACGCCTTCCTTCTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG 315

Qy    301 AAATTTTTGTGAGGCTGTCCATATCATCATCACACTGTCAACCTCTCACAGCAGCGTTCTC 360
          || ||| |||||||||||| ||||||||||||||||||||||||  ||||||||||||||||
Db    316 AACTTCCTATGCAAGGCAGTCCATGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTC 375

Qy    361 ATCCTGGCCTTCATCAGCCTGACCGTACCTGCGCATTGTCCACGGCCACCAACAGTCAA 420
          ||||||||||||||||||||||||  |||||||||||||||  |||||||||||||||||
Db    376 ATCCTGGCCTTCATCAGTCTGGACCGTCACCTGCCATCGCTACCGCCACCAACAGTCAG 435

Qy    421 AGCCCAAGGAAACGTCTGGCTGGAAAAGCAGTCTATGTGGGGTCGGGCGTCTCGAGCCCAGCCTC 480
          ||||||||||||| ||||||||||||||||||||||||||| |||||| | ||||||||
Db    436 AGCCCAAGGAAGCTGTTGGCTGGAAAAGGTGGTCTATGTTGGGGTCTGGATCCCTGCCCTC 495

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCAGCGGGACACATCAGTCAGGGG 540
          ||  || |  |||||||||||||||||||||
Db    496 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG----------TCAGTGAGGCA 540

Qy    541 GATGACAGGTACATCTGTGACCGGCCTTACCCCGATAGCCTGTGGGATGGTGGTGTTTCAA 600
          |||||||| ||||||||||||| |||||| |||| ||||||| |||| ||||||| ||||||
Db    541 GATGACAGATATCTGTGACCGGCCTTCACCCCCAATGACTGTGGGTGTGTGTTCCAG 600

Qy    601 TTCCAGCATAATGGGTGGCTGTCATCATCCGGCATCGGCGATCCTCCTCCTGTTACTGC 660
          ||||| | | ||| ||| |||| ||||||||| ||| |||||| ||| ||||||||||||||
Db    601 TTTCAGCACATCATGGTTGGCCTTATCATCTGGCTGGTATTGTCATCCTGCTATTGC 660

Qy    661 ATCATCATCCTCTAAGCTGTCACACTCCAAGGGCCAGGCCCTCAAGACG 720
          ||||||||| |||||| |||| | |||| ||||| |||| ||||||||
Db    661 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAGGCCCTCAAGACC 720

Qy    721 ACAGTCATCCTCATCCTCAGCTTTCTTGCCTGCTGGTGCCATATATGTGGGATCAGC 780
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    721 ACAGTCATCCTCATCCTCGGCTTTCTTCGCTGTGGTGCCTTACACATTGGGATCAGC 780

Qy    781 ATCGACTCCTTCATCCTTTTGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
          |||||||||||||||||||||  |||||||| |||||||||||||| || |||||||||
Db    781 ATCGACTCCTTCATCCTTCTGGAAATCATCAGGCAAGGTGTGAGTTTGAGAACACTGTG 840

Qy    841 CACAAGTGGATCTCCATCACAGAGGCCCTGCCTTCCACTGTGCCTGAACCCCATC 900
          |||||||||| ||||||||||||||||||||||| |||| |||  ||||||||||||
Db    841 CACAAGTGGATTTCCATCACCGAGGGCCCTAGTTTCTTCCACTGTTGTGAACCCCATC 900

Qy    901 CTTCTATGCCTTCCTCCGGGGCCAAGTTCAAAAGCTTCGCCCAGCATGCACTCAACTCCATG 960
          ||||| || |||||||| |||| |||| || |||| ||| |
Db    901 CTCTATGCTTTCCTTGGAGCCAAATTTAAAACCTTCCGCCAGGCACCACTACCTCTGTG 960

Qy    961 AGCAGAGGGTCCAGCCTCAAGATCCTTCCAAAGGAAGATCCTTCCGTC 1020
          ||||||||| ||||||||  ||||||  ||||||||||| ||||||||||
Db    961 AGCAGAGGGTCCAGCCTCCAGATCCTCCAAAGGAAGAAGCAGGTGCACATTCATCGTT 1020

Db   1064 GACTTTTTTTATACGATAAATAACTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1123

Qy   1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTAATTGACTGGTGGGAGTTTATGTTCC 1200
          |||||||| | ||||||||||||||||||||||| || ||||||||||||| ||||||
Db   1124 ACTGACCAATATTGTACAGTTTTTATTGCTTGTTGGATTTTGT------CTTGTGTTTC 1177

Qy   1201 TCTAGTTTTTGTGAGGTTGACTAATTTATATAAAATATTGTTTTTGTTTTGTTTCATGT 1260
          ||  ||||| ||||||  |||||||||||| |||||||||||| ||||||||||||
Db   1178 TTTAGTTTTTGTGAAGTTTAATTGACTTATTTATA---TAAATTTTTTTGTTTTCATAT 1233

Qy   1261 GAATGAGCCTCTAGGCAGGACCTGGGCCAAGTTCTTAGTAGCTGTTATCTGTGTGTAG 1320
          ||  ||||||||||||||||||||||||||||||||| |||||| |||  |||||||||
Db   1234 TGATGTGTCTCAGGCAGGACCTGGGCCAAGTTCTTAGTTGCTGTATGTCTCGTGGTAG 1293

Qy   1321 GACTGTAGAACTGTAGAGGAGAAGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
          |||||||||        ||||| || || |||||||||||  ||||||||||
Db   1294 GACTGTAGAA-------AAGGGACTGAACAATTCCAGAGAGGCGTGTAGTGAATCACGTAAA 1345

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
          ||||| ||||| |||| ||||||||| ||||
Db   1346 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1376
```

RESULT 4

AA227610

ID   AA227610 standard; DNA; 1588 BP.
XX
AC   AA227610;
XX
DT   16-DEC-1999   (first entry)
XX
DE   Human CXCR4 coding sequence.
XX
KW   CXCR4; human; neovascularisation; inhibitor; solid cancer; therapy;
KW   tissue repairing agent; vascularisation; ss.
XX
OS   Homo sapiens.
XX
PN   WO9948528-A1.
XX
PD   30-SEP-1999.
XX
PF   23-MAR-1999;   99WO-JP001448.
XX
PR   24-MAR-1998;   98JP-00095448.
XX
PA   (CHUS) CHUGAI SEIYAKU KK.
PA   (KISH/) KISHIMOTO T.
XX
PI   Kishimoto T, Nagasawa T, Tachibana K;
XX
DR   WPI; 1999-591042/50.
DR   P-PSDB; AAY39993.
XX
PT   CXCR4-potentiating agents and methods useful for inhibiting
PT   neovascularization, and treating solid cancers.
XX
PS   Disclosure; Page 48-49; 63pp; Japanese.
XX
CC   This sequence encodes the human CXCR4 protein. The invention relates to
CC   remedies inhibiting neovascularisation, remedies for solid cancer,
CC   remedies for diseases pathologically caused by neovascularisation and
CC   tissue repairing agents containing as the active ingredient a substance
CC   capable of potentiating CXCR4. Based on a finding that vascularisation is
CC   inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC   remedies inhibiting vascularisation which contain as the active
CC   ingredient a substance capable of potentiating CXCR4, remedies for solid
CC   cancer, remedies for diseases pathologically caused by neovascularisation
CC   and tissue repairing agents containing as the active ingredient a
CC   substance capable of potentiating CXCR4. It is also possible to establish
CC   methods for treatment with the use of these remedies
XX
SQ   Sequence 1588 BP; 396 A; 375 C; 334 G; 483 T; 0 U; 0 Other;

Query Match
Best Local Similarity 55.1%; Score 933.4; DB 3; Length 1664;
81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

```
Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG   60
Db  92 ATATACACTTCAGATAACTACACCGAGGAAATGGCTCAGGGACTATGACTCCATGAAG  151

Qy  61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC  120
Db 152 GAACCCTGTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC  211

Qy 121 TTCATCATCTTCTTGACTGGCAATGTCGGCAATGTGGGCAATGGGTGTCATCCTGGTTAC 180
Db 212 TCCATCATCTTCTTAACTGGCAATTGTGGGCAATGGTTGGTCATCCTGGTCATGGGTTAC 271

Qy 181 CAGAAGAGTCAAGGAGCATGAGGACAAGAAGTACCGGCTGCAGTGGTCGACCCTC      240
Db 272 CAGAAGAAACTGAGAAGCATGACGGACAAGAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC 331

Qy 241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCAGTGCTGATGGCTACTTTGGG 300
Db 332 CTCTTTGTCATCACGCTCCCTTCCCTTCTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG 391

Qy 301 AAATTTTGTGTAAGGCTGTCCATATCATCTGTGGGCAATGGATTGGTCATCCTGTCTCTC 360
Db 392 AACTTCCTATGCAAGGCAGTCCATGTCATCGTCATCTACACAGTCAACCTCTACAGTGTCCTC 451

Qy 361 ATCCTGGCCTTCATCAGCTGGACCGGTACCTGCTACCTGCTCCATTGTCCACGGCCACCAACAGTCAA 420
Db 452 ATCCTGGCCTTCATCAGTCTGGACCGGTACCTGGCCATCTGGCCGGTCTATGTGGGCGTTCATGTCCTCC 511

Qy 421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTTGGGCGTCTGGGAGTCCTGCCCCTC 480
Db 512 AGGCCAAGGAAGGAGCGTTGCGTGAGTCTGAAAAGGTGGTCTATGTTGGGCGTCTGGGGATCCCGCCCCTC 571

Qy 481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTGCAGCATCAGTCAGGGG 540
Db 572 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG---------TCAGTGAGGGGCA 616

Qy 541 GATGACAGGTACATCTGTGACCGGCTTTACCCGATAGCCTGTGGATGGTGGTGTTCAA 600
Db 617 GATGACAGATATCTGTGACCGGCTTCTCACCCCAATAGACTTGTGGGTGGTTGTTCCAG 676

Qy 601 TTCCAGCATATAATGGTGGGTCCATCCTCGTCATCCTGTCCTGTTACTGC 660
Db 677 TTTCAGCACATCATGGTGGCCTTATCCTGCTGTATTGTCATCCTGCTATTGC 736

Qy 661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACG 720
Db 737 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGGAGCGCAAGGCCCTCAAGACC 796

Qy 721 ACAGTCATCCTCATCCTAGCTTTCTTGCTCTGCTGGCTGCCATATATGTGGGATCAGC 780
Db 797 ACAGTCATCCTCATCCTCGCTTTCTTCTCCTGTTGGCTGCCTTACTACATTGGGATCAGC 856

Qy 781 ATCGACTCCTTCATCCTTTGGGAGTCATCAGCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
```

Qy 1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCAGCCTTATGCAAAGACTTATAT 1080
Db 1021 TCCACTGAGTCTTGAGTCTTCAAGTTTTCACTCCAGCTAA------CACAGATGTAAAA 1072

Qy 1081 AATATATAATATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAAGAG 1140
Db 1073 GACTTTTTTTTATACGATAAATAACTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1132

Qy 1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTAATTGACTGTGGGAGTTTATGTTCC 1200
Db 1133 ACTGACCAAATATGTACAGTTTTTATTGCTGTGTGGATTTTGT-----CTTGTGTTTC 1186

Qy 1201 TCTAGTTTTTGTGAGGTTGACTTAATTTATATAAAATATTGTTTTTGTTGTTCATGT 1260
Db 1187 TTTAGTTTTTGTGAAGTTAATTGACTTATTTATA----TAAATTTTTTTGTTTCATAT 1242

Qy 1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
Db 1243 TGATGTGTGTCTAGGCAGGACCTGTGCGCCAAGTTCTTAGTTGCTGTGTAGTGTCGTGGTAG 1302

Qy 1321 GACTGTAGAACTGTAGAGGAGGAAGAAACTGACAATTCCAGAGATGTGTGGTAAATTGAATAAA 1380
Db 1303 GACTGTAGAA-------AAGGAACTGACAATTCCAGAGACGTGTAGTTAATCACGTAAA 1354

Qy 1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
Db 1355 GCTAGAAATGATCCCCAGCGTGTTTATGCATA 1385

PT   expression.
XX
PS   Disclosure; Page 5-6; 6pp; Japanese.
XX
CC   This sequence represents a CXCR4 coding sequence. The invention relates
CC   to antisense oligonucleotides that hybridise specifically with
CC   chromosomal DNA and/or RNA encoding a CXCR4 protein, and inhibit the
CC   expression of the CXCR4 protein. The antisense oligonucleotides can be
CC   used in an anti-HIV agent. The oligonucleotides and the anti-HIV agent
CC   are highly effective as a preventive and treating agent for HIV infection
XX
SQ   Sequence 1664 BP; 414 A; 398 C; 359 G; 493 T; 0 U; 0 Other;

Query Match   55.1%;   Score 933.4; DB 3;   Length 1664;
Best Local Similarity   81.8%;   Pred. No. 4.3e-207;
Matches 1154;   Conservative 0;   Mismatches 216;   Indels 41;   Gaps 5;

```
Qy     1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG   60
         ||||||||||| ||||||| |||   ||||||   ||| || || |||||||||||  |||
Db    92 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGACTATGACTCCATGAAG  151

Qy    61 GAACCCTGCTTCCGGGATGAAAAGTCCATTCAATAGGATCTCCTGCCACCATCTAC      120
         ||||||| ||||| |||| |||||| |||||| |||||| |||||| || ||||||
Db   152 GAACCCTGTTTCCGTGACGAAGAAATGCTAATTCAATAAAATCTTCTGCCACCATCTAC   211

Qy   121 TTCATCATCTTCTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC   180
         |||||||||||| || ||||||| |||||||| |||||| |||||||||||||||||||
Db   212 TCCATCATCTTCTTAACTGGCATTGTGGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC   271

Qy   181 CAGAAGAAGCTAAGGACCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC   240
         ||||||||| |||||||||||| |||||||||||||||| |||||||||||||||||||
Db   272 CAGAAGAAACTGAGAGACGATGACCGGCTGCACCTGTCACCTGTGCACCTGCCGACCTC   331

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCCAGTTGACTGTACTTTGG   300
         |||||| ||||| ||||||||||||||||||| || |||| || |||||| |||||||||
Db   332 CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCATGGCCAAATGGTACTTTTGGG   391

Qy   301 AAAATTTTGTGTAAGCTGTCCATATCATCACTGTCACACACAGCAGTGTCCTC      360
         ||| |||| ||||| ||||| ||||||||| || |||||||||||||||||
Db   392 AACTTCCTATGCAAGGCAGTCCATCATCTCACCGTCACACACAGCAGTGTCCTC   451

Qy   361 ATCGGCCTTCATCAGCTGGCTGAAAAGGCAGTCTATGTGGCGTCTGGATCCCAGCCCTC   420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   452 ATCGGCCTTCATCAGCTGGCTGAAAAGGCAGTCTATGTGGCGTCTGGATCCCGCCCCTC   511

Qy   421 AGGCCAAGGAAACTCTGCTGAAAAGGCAGTCTATGTGGCGTCTGGACATCAGTCAGGGG   480
         |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   512 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTGGCGTCTGTATGTCAGGGG   571

Qy   481 CTCCTGACTATACCTGACTTCATCTTGCCGAGTCAGCGAGGGACATCAGTCAGGGG      540
         |||||| |||| |||||||||||| || |||||   ----------    ||||||
Db   572 CTGCTGACTATTCCCGACTTCATCTTGCCAACG----------TCAGTGAGGCA   616

Qy   541 GATGACAGGTACATCTGTGACGCGGCCTTTACCCCGATAGCCTGTGATGGTGGTTTCAA   600
         |||| |||||||||||||||| ||||| |||||||| |||||| |||||||||| |||
Db   617 GATGACAGATATATCTGTGACCGGCTTCACCCCAATGACTTGTGGTGGTGTTGTTCAG   676

Qy   601 TTCCAGCATATAANGTGGGTCATCCTCGCCGGCATCGTCATCCTCCTGTTACTGC      660
         ||||| |||| |||| |||||||||||| |||| ||||||||| |||||||||||
Db   677 TTTCAGCACATCATGTTGGCCTTATCCTCGCCTGTATTGTCATCCTGCTATTGC   736

Qy   661 ATCATCTTCTAAGCTGTCACACTCCAAGGGCCACCAGGGCCCTCAAGACG   720
         |||||||| |||||||||||||||||||||||||||||||||| |||||
Db   737 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCCTCAAGACC   796

Qy   721 ACAGTCGTCATCCATCCATCTAGCTTTCTTCTTGCTGCTGCCATATTATGTGGGGATCAGC   780
         ||||| |||||||||||||||| || ||||||||||||||||| |||||||||||||||
Db   797 ACAGTCATCCATCCATCCATCTGGCTTTCCTTGGTGCTGTGCCTTACTACAATGGGATCAGC   856

Qy   781 ATCGGACTCCCTTCATCCTCCGTGAAAATCATCAAGCAAGGGGTGTGAGTTTGAGAACTGTG   840
         ||||||||||||||||||||||||||| |||||| ||||||||| |||||||||||||||
Db   857 ATCGGACTCCCTTCATCCTCCGTGAAAATCATCAAGCAAGGGGTGTGAGTTTGAGAACACTGTG   916

Qy   841 CACAAGTGGATCTCCATCACAGAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC   900
         ||||||||||| ||||||||| ||||||||| |||||||||||||||| |||||||||||
Db   917 CACAAGTGGATTTCCATCACCGAGGCCCTAGCTTCTTCCACTGTTGTCTGAACCCCATC   976

Qy   901 CTCTATGCCTTCCTCGGGCCAAGTTCAAAAGCTCTGCCCCAGCATGCACTTCAACTCCATG   960
         |||||||| ||||| || |||||  ||||| ||||||||||||| ||||| ||||||||||
Db   977 CTCTATGCTTTCCTTGGAGCCAAATTTAAAACCTCTGCCCCAGCACGCACCTCACCTCTGTG   1036

Qy   961 AGCAGAGGCTCCAGCTTCAAGATCCTTTCCAAAGAAGCGGGGTGGACACTCTTCCGTC   1020
         |||||||| ||| |||||||||||||| |||||||| ||| || ||||| |||| |||
Db  1037 AGCAGAGGGTCCAAGCCTCCAAGATCCTCAAGATCCTCCCAAAGGAAAGCGAGGTGGACATTCATCTGTT   1096

Qy  1021 TCCCAGCGAGTCAGAATCCTCCAGTTTCACTCCAGTTTTCACTCCAGCTAA   1080
         |||||| |||||||| ||| ||||||||||||||||| ||||||||||||
Db  1097 TCCCACTGAGTCTGAGTCTTCAGTGTTCACTCCAGTTTTCACTCCAGCTAA------CACAGATGTAAAA   1148

Qy  1081 AATATATATATATATGATATAAGAACTTTTTTATGTTACACATTTCCAGATATAAAGAG   1140
         ||| || || ||| ||||| ||||||||||||||||| ||||| |||||||||||
Db  1149 GACTTTTTTTTTATACGATAAATAACTTTTTTTAAGTTACACATTTTTCAGATATAAAAAG   1208

Qy  1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTTTATTGACTGTTGGAGTTTATGTTCC   1200
         |||||| |||||||| |||||||||||||||||| |||||||||||||| ||| ||||||
Db  1209 ACTGACCAATATTGTACAGTTTTTATTGCTTGTGTTGGATTTTTGT-----CTTGTGTTTC   1262

Qy  1201 TCTAGTTTTGTTGTGAGGTTGACTTAATTATATATTATATAAATATTGTTTTTTTGTTCATGT   1260
         |||||||||||||||||||||| |||||||| |||||||| || || ||||||||||||||
Db  1263 TTTAGTTTTGTGAAGTTAATTGACTTATTTATT-----TAAATTTTTTTTGTTTTCATAT   1318

Qy  1261 GAATGAGCGCTAGGCAGGACCCTGTGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG   1320
         |||||||| ||||| ||| |||||||||||||| || ||||| |||||||||||||||
Db  1319 TGATGTGTGTCAGGCAGGACCCTGGCCAAGTTCTTAGTTGCTGTGTGATGTCTCGTGGTAG   1378

Qy  1321 GACTGCTAGAACGTGTAGAGGAGAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAA   1380
         |||| ||| ||   ----- ||| |||||||||| |||||| |||||| ||||||||||
Db  1379 GACTGTAGAA--------AAGGGAACTGAACATTCCAGAGTGTGTGAGTTAATTCACGTAAA   1430

Qy  1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA   1411
         ||||||| ||| ||||||||||||| |||||
Db  1431 GCTAGAAATGATCCCCAGCTGTTTATGCATA   1461
```

RESULT 6
AAZ40014
ID   AAZ40014 standard; DNA; 1664 BP.
XX
AC   AAZ40014;
XX
DT   14-FEB-2000 (first entry)
XX
DE   CXCR4 coding sequence.
XX
KW   CXCR4; inhibitor; antisense oligonucleotide; anti-HIV agent;
KW   HIV infection; therapy; ss.
XX
OS   Synthetic.
XX
PN   JP11285391-A.
XX
PD   19-OCT-1999.
XX
PF   18-NOV-1998;   98JP-00327942.
XX
PR   19-NOV-1997;   97JP-00335085.
XX
PA   (HISM) HISAMITSU PHARM CO LTD.
PA   (TOAG) TOA GOSEI CHEM IND LTD.
XX
DR   WPI; 2000-026817/03.
XX
PT   An oligonucleotide used as an anti-HIV agent - inhibits CXCR4 protein

PA　(AMLI/) AMLING M.
XX
PI　Amling M, Karsenty G, Ducy P;
XX
DR　WPI; 2001-488709/53.
DR　P-PSDB; AAE06690.
XX
PT　Treating or preventing reduced bone mass, e.g. osteoporosis, by reducing
PT　the level of neuropeptide Y activity in blood or cerebrospinal fluid.
XX
PS　Example 7; Page 84-85; 102pp; English.
XX
CC　The present invention relates to a method for treatment or prevention of
CC　bone diseases characterised by loss of bone mass, comprises administering
CC　to a mammal a compound that lowers the level of neuropeptide Y (NPY) in
CC　the serum or cerebrospinal fluid (CSF) or a compound that lowers the
CC　level of inositol phosphate (IP) or extracellular signal-regulated kinase
CC　(ERK). The method is specifically used to treat (including by gene
CC　therapy) or prevent osteoporosis, osteopenia or Paget's disease, but may
CC　also be used e.g., in cases of fractures or bone metastases. These
CC　diseases may also be diagnosed by detecting elevated NPY levels,
CC　including monitoring of treatment, assessing efficacy of compounds in
CC　clinical trials and for identifying subjects at risk. The present
CC　sequence is a human NPY Y3 receptor DNA
XX
SQ　Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match　　55.1%;　Score 933.4;　DB 4;　Length 1679;
Best Local Similarity　81.8%;　Pred. No. 4.3e-207;
Matches 1154;　Conservative　0;　Mismatches 216;　Indels 41;　Gaps 5;

Qy　　1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGTCTGAGACCATGACTCCAACAAG 60
　　　　|||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db　104 ATATACACTTCAGATAACTACTACACCGAGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163

Qy　 61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
　　　　||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||||
Db　164 GAACCCTGTTCCGTGGATGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC 223

Qy　121 TTCATCATCTTCTTGACTGGCATAGTGGCATTGTGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
　　　　|||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db　224 TCCATCATCTTCTTAACTGGCATTGTGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 283

Qy　181 CAGAAGAAGCTAAGGAGCCATGACGGACAAGTACCGGTCGCACCTGTCAGTGGCTGACCTC 240
　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　284 CAGAAGAAACTGAGAAGCCATGACGGACAAGTACAGGTCGCACCTGTCAGTGGCCGACCTC 343

Qy　241 CTCTTGTCATCACACTCCCCTTCGGGCAGTTGATGCCATGGCTGACGTACTTTGGG 300
　　　　||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db　344 CTCTTTGTCATCATCGCTTCCCTTCTGGGCAGTTGATGCCGTGGCAAACTGTACTTTGGG 403

Qy　301 AAATTTTGTGTAAGGCTGTCCATATCATCTCACACTGTCAACCTCACAGACAGCGTCTC 360
　　　　||||||||||||| |||||||||| |||||||||||||||||||| |||||||||||||
Db　404 AACTTCCTATGCAAGGCAGTCCATGTCATCTCACACAGTCAACCTCTCACAGCAGTGTCCTC 463

Qy　361 ATCCTGGCCTTCATCAGCCTGGACGGTACCTGCCATTGTCCAGCGCCACCAACAGTCAA 420
　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　464 ATCCTGGCCTTCATCAGTCTGGACCGCTGGCCATCATCGTCCACGCGCCACCAACAGTCAG 523

Qy　421 AGGCCAAGGAAACTGCTGGCTGAAAAGCAGTCTATGTGGGGTCTATGTTGGGCTGGCCCTC 480
　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　524 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTTGGGCTCTGGATCCCTGCCCTC 583

Qy　481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGAGCCTTCAGCAGTCAGGGG 540
　　　　|||||||||||||||||||||||||||||| |||||||||||||| ||||||||
Db　584 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG-----------TCAGTGAGGGCA 628

Qy　541 GATGACAGTACATCTGTGACCGCCTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA 600
　　　　|||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||
Db　629 GATGACAGATATATCTGTGACCGGTTCTACACCCCAATGACTTGTGGGGTGGTGTTCCAG 688

Qy　601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTTGTTACTGC 660
　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy　841 CACAAGTGGATCTCCATCACAGAGAGGCCCTGCCTTCCACTGTGCCTGACCCCATC 900
　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　917 CACAAGTGGATTCATCACCGAGGGCCCAGCTTTCTTCCACTGTTGCTGAACCCCATC 976

Qy　901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCGCCCAGCATGACTCACTTCCATG 960
　　　　||||||||||||||||||||| |||||||||| ||| |||||
Db　977 CTCTATGCCTTCCTTGGAGCCAAATTAAAACCTGCCCAGCACGCACGCACCTCTGTG 1036

Qy　961 AGCAGGAGGCTCCAGCCTCAAGATCCTTCAAAGGGAAAGCGGGTGGACACTCTTCCGTC 1020
　　　　||||||||||||||||| |||||||||||| ||||||||||||||||||||||| ||| ||
Db　1037 AGCAGGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTT 1096

Qy　1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
　　　　||||||||||||||||||||| |||||||| |||||||||||||| |||| ||||
Db　1097 TCCACTGGTCTGAGTCTTCAAGTTTCACTCCCAGCTAA------CACAGATGTAAAA 1148

Qy　1081 AATATATATATATATGATAAAGAACTTTTTTATGTTACACATTTCCAGATATAAGAG 1140
　　　　||||||||||||||||||||||| |||||||| |||||||||| |||||||||||
Db　1149 GACTTTTTTTTTATACGATAAAATAAACTTTTTTTAGTTACACATTTTCAGATATAAAAG 1208

Qy　1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTAATTGACTGGTGGGAGTTTATGTTCC 1200
　　　　||||||||||||||| |||| |||||||||| |||||| ||||||||||||||||||
Db　1209 ACTGACCAAATTGTACAGTTTTATTGCTTGTTGGATTTTTGT-------CTTGTGTTTC 1262

Qy　1201 TCTAGTTTTTGTGAGGTTGACTTAATTATATAAATATTGTTTTTTGTTTTCATGT 1260
　　　　||| |||||||||||||| |||||||||||||||||||||| |||||||||||||||
Db　1263 TTTAGTTTTTGTGAAGTTGAAGTTTAATTGACTTATTTATA----TAAATTTTTTTTTGTTTTCATAT 1318

Qy　1261 GAATGAGCGTCTAGGCAGGACCTGTTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
　　　　||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||||
Db　1319 TGATGTGTGTCTAGGCAGCAGACCTGTGGCCAAGTTCTTAGTTGCTGTGTATGTCTCGTGGTAG 1378

Qy　1321 GACTGTAGAACTGTAGAGGAGAAACTGAACATTCCAGAAATGTGGTAAATTGAATAAA 1380
　　　　|||||||||　　||||||||
Db　1379 GACTGTAGAA------AAGGGACTGAAACATTCCAGAACGCGTGTAGTTAATCACGTAAA 1430

Qy　1381 GCTAGCCGTGATCCTCCAGCTGTTGCTGCATA 1411
　　　　||||| ||||||||||||||
Db　1431 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1461

RESULT 7
AAD12801
ID　AAD12801 standard; DNA; 1679 BP.
XX
AC　AAD12801;
XX
DT　16-OCT-2001 (first entry)
XX
DE　Human neuropeptide Y (NPY) Y3 receptor DNA.
XX
KW　Human; neuropeptide Y; NPY; bone disease; bone mass; gene therapy;
KW　cerebrospinal fluid; CSF; inositol phosphate; IP; Paget's disease;
KW　fracture; extracellular signal-regulated kinase; ERK; osteoporosis;
KW　osteopenia; bone metastasis; neurotransmitter; osteogenic;
KW　NPY Y3 receptor; ds.
XX
OS　Homo sapiens.
XX
FH　Key　　Location/Qualifiers
FT　CDS　　89..1147
FT　　　　/tag= a
FT　　　　/product= "Human neuropeptide Y (NPY) Y3 receptor"
XX
PN　WO200153477-A1.
XX
PD　26-JUL-2001.
XX
PF　22-JAN-2001; 2001WO-US002040.
XX
PR　20-JAN-2000; 2000US-00489872.
XX
PA　(BAYU) BAYLOR COLLEGE MEDICINE.

XX
PF 28-FEB-2001; 2001WO-US006322.
XX
PR 02-MAR-2000; 2000US-00517605.
XX
PA (UYNY) UNIV NEW YORK STATE.
PA (UYNI-) UNIV NIJMEGEN.
XX
PI Littman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;
XX
DR WPI; 2001-602565/68.
XX
PT An antibody for the treatment or prevention of HIV-infection comprises a
PT gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of
PT DC-SIGN due to concomitant conformational change.
XX
PS Disclosure; Page 126-127; 131pp; English.
XX
CC The specification describes an antibody which is specific for an
CC antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or
CC is exposed upon gp120 binding of DC-SIGN due to concomitant
CC conformational change. DC-SIGN is a receptor that is specifically
CC expressed on dendritic cells and facilitates infection of T lymphocytes
CC with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC
CC -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)
CC with high affinity. The antibody of the invention inhibits the trans
CC enhancement of HIV entry into a T cell or macrophage facilitated by
CC dendritic cells. The antibody is useful to treat or prevent HIV
CC infection. The present sequence represents a human polynucleotide, which
CC is used in the course of the invention
XX
SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   55.1%; Score 933.4; DB 4; Length 1679;
Best Local Similarity 81.8%; Pred. No. 4.3e-207;
Matches 1154; Mismatches 215; Conservative 0; Indels 41; Gaps 5;

```
Qy    1 ATATACACTTCTGATAACTACACTCTGAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
         |||||||||||  ||||||| |||||||  |||||||  ||||||||||  ||  |||||||
Db  104 ATATACACTTCAGATAACTACACCGAGGAAAATGGGCTCAGGGACTATGACTCCATGAAG 163

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCCATCCTAC 120
         ||||||||||  |||||||||| ||||||||||||||||   |||||||||||||||||
Db  164 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAATCTTCCTGCCCACCATCTAC 223

Qy  121 TTCATCATCTTCTTGACTGCATAGTCGGCAATGAGTTCGGTCGTGTCATGGGTTAC 180
         ||||||||||  |||||  || |||| ||||||||||||||||||||||||||||||
Db  224 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGTCGTGTCATGGGTTAC 283

Qy  181 CAGAAGAAGCTAAGGAGCATGACGGACAAGAAGTACCGGCTGCACCTGTCAGTGGCCCTC 240
         ||||||||||  ||||||||||||||||||||||||  ||||||||||  ||||||||||
Db  284 CAGAAGAAACTGAGAAGCATGACGGACGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC 343

Qy  241 CTCTTTGTCATCACACTCCCCTTCCCTTCGGGCAGTGGTGATGCCGAGTGATGGCCTGG 300
         ||  |||||||||||||||||  |||||||||||||||||  |||||||||||||||||
Db  344 CTCTTTGTCATCGCGCTTCCCTTCCCTTCGGGCAGTGGCGTGGCAAACTGGTACTTTGGG 403

Qy  301 AAATTTTTGTGTAAGGCTGTGTCCATATCATCTACACTCTACAGCAGCGGTTCTC 360
         |||||||||||  ||||||| ||  |||||  |||| ||||||||||||||||||
Db  404 AACTCCATGCAAGGCAGTCATGTCATCTCACACAGTCAACCTCTACAGCAGCGTCCTC 463

Qy  361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCCCCATTGTCCACGCCACCAACAGTCAA 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  464 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCCCCATTGTCCACGCCACCAACAGTCAG 523

Qy  421 AGGCCAAGAAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCGATCCCTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  524 AGGCCAAGGAAGCTGCTGGCTGAAAAGGGTCATTGTTGGCGTCTGGATCCCGCCCTC 583

Qy  481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGTGGGGACATCAGTCAGGGG 540
         |||||||  |||  ||||||||||||||  ||||||||||||||||||||||||
Db  584 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG------------TCAGTGAGGCA 628
```

```
Db  689 TTTCAGCACATCATGGTTGGCCTTATCCTGCCTGGTATGTCATCCTGCCTGCTATTGC 748

Qy  661 ATCATCATCCTAAGCTGTCACACTCCAAGGGCCAACCAGCAAGGGCCAAGACG 720
Db  749 ATTATCATCCCAAGCTGTCACACTCCAAGGGCCACACTCCAAGGGCCAAGACC 808

Qy  721 ACAGTCATCCTCATCCTAGCTTCTTTGCCTGGCTGCCATTATTGTGGGGATCAGC 780
Db  809 ACAGTCATCCATCCTCTGGCTTTCTTCGCCTGTTGGCTGCCTTACTACATTGGGATCAGC 868

Qy  781 ATCGACTCCTTCATCCTTTTGGGAGTGCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
Db  869 ATCGACTCCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 928

Qy  841 CACAAGTGGATCCATCACAGAGGCCCGCCTTCTTGCCTGTTGCTGCCTGAACCCCATC 900
Db  929 CACAAGTGGGATTTCCATCACCAGGCCCTAGCTTTCTTCTTCCACTGTTGCTGAACCCCATC 988

Qy  901 CTCTATGCCTTCCTCCTGGGGCCAAGTTCAAAAGCTGCTGCCAGCATGCACTCAACTCCATG 960
Db  989 CTCTATGCTTTCCTTGGAGCCAAATTTAAAACCTGCCCAGCACGCACCACTCTGTG 1048

Qy  961 AGCAGAGGTCCAGCCTCAAGATCCTTCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
Db 1049 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTT 1108

Qy 1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTAACCCTATGCAAAGACTTATAT 1080
Db 1109 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAA-------CACAGATGTAAAA 1160

Qy 1081 AATATATAATATATATGATAAGAACTTTTTTATGTTACACATTTTTCCAGATATAAGAG 1140
Db 1161 GACTTTTTTTTTATACGATAAATAACTTTTTTTTAAGTTACACATTTTTCAGATATAAAAG 1220

Qy 1141 ACTGACCAGTTGTACAGTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
Db 1221 ACTGACCAAATATTGTACAGTTTTTTATTGCTTGTCTGTGGATTTTTGT------CTTGTGTTTC 1274

Qy 1201 TCTAGTTTTGTGAGGTTTGACTTAATTTATATAAATATTGTTTTTTGTTTCATGT 1260
Db 1275 TTTAGTTTTTTGTGAGGTTTAATTGAAGTTTAATTATA----TAAATTTTTTTGTTTCATAT 1330

Qy 1261 GAATGAGCGTCTAGGCAGGAACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
Db 1331 TGATGTGTGTGTAGGCAGGACCGTGTGGCCCAAGTTCTTTGTCTCTATCTGTGGTAG 1390

Qy 1321 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
Db 1391 GACTGTAGAA-------AAGGGACTGAACATTCCAGAGACGTGTAGTGAATCACGTAAA 1442

Qy 1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
Db 1443 GCTAGAAAATGATCCCCAGCTGTTTATGCATA 1473
```

RESULT 8
AAI65467
ID AAI65467 standard; DNA; 1679 BP.
XX
AC AAI65467;
XX
DT 10-DEC-2001 (first entry)
XX
DE Nucleotide sequence of a human polynucleotide.
XX
KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;
KW C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.
XX
OS Homo sapiens.
XX
PN WO200164752-A2.
XX
PD 07-SEP-2001.

KW restenosis; ischaemic stroke; spinal cord injury; infection; ulcer;
KW human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;
KW AIDS; pain; cancer; benign prostatic hypertrophy; diabetes; obesity;
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW hypotension; hypertension; urinary retention; osteoporosis; stroke;
KW anginal pectoris; myocardial infarction; benign prostatic hypertrophy;
KW migraine; vomiting; psychotic disorder; neurological disorder; anxiety;
KW schizophrenia; dyskinesia; Huntingdon's disease; restenosis;
KW inflammatory disease; rheumatoid arthritis; diabetic retinopathy;
KW inflammatory bowel disease; atherosclerosis; Alzheimer's disease;
KW congestive heart failure; cardiac remodeling; angiogenic diseases;
KW solid tumour; Kaposi Sarcoma; human; gene; ss.
XX
OS Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT CDS   89..1147
FT   /*tag= a
FT   /product= "CXCR4"
FT   /note= "CXC chemokine receptor 4"
XX
PN US2002107195-A1.
XX
PD 08-AUG-2002.
XX
PF 17-SEP-2001; 2001US-00953692.
XX
PR 21-JUL-1998; 98US-0093596P.
PR 21-JUL-1999; 99US-00358624.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Gupta SK;
XX
DR WPI; 2002-697879/75.
DR P-PSDB; ABG32977.
XX
PT Inducing chemotaxis of endothelial cells, useful for regulating
PT angiogenesis, e.g. for treating cancer, comprises treatment with stromal
PT cell-derived factor 1 alpha.
XX
PS Disclosure; Fig 1; 26pp; English.
XX
CC The invention describes a method of inducing chemotaxis of endothelial
CC cells by treatment with stromal cell-derived factor 1alpha (SDF1a). The
CC method is used for stimulating EC chemotaxis, and thus angiogenesis, and
CC is used for treating atherosclerosis, restenosis, ischaemic stroke and
CC spinal cord injury. Inhibition of this process is useful in treatment and
CC prevention of a very wide range of diseases, such as, infection (by
CC bacteria, fungi, protozoa or viruses such as human immunodeficiency virus
CC (HIV) and acquired immunodeficiency syndrome (AIDS)), pain, cancer and
CC benign prostatic hypertrophy, diabetes, obesity, anorexia, bulimia,
CC asthma, Parkinson's disease, acute heart failure, hypotension,
CC hypertension, urinary retention, osteoporosis, anginal pectoris,
CC myocardial infarction, stroke, ulcers, benign prostatic hypertrophy,
CC migraine, vomiting, psychotic and neurological disorders (e.g. anxiety,
CC schizophrenia) and dyskinesias (e.g. Huntingdon's disease), inflammatory
CC disease, rheumatoid arthritis, diabetic retinopathy, inflammatory bowel
CC disease, atherosclerosis, restenosis, Alzheimer's disease, congestive
CC heart failure, cardiac remodeling, angiogenic diseases, solid tumours,
CC and Kaposi Sarcoma. This sequence encodes the human CXC chemokine
CC receptor 4 (CXCR4) polynucleotide
XX
SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   55.1%;   Score 933.4;   DB 6;   Length 1679;
Best Local Similarity   81.8%;   Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy   1 ATATACACTTCTGATAACTACACTCTGGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
         |||||||||||||||| |||||||| ||||||||| ||||||||| | |||||||||||||
Db   104 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163

Qy   541 GATGACAGGTACATCTGTGACCGCCTTACCCGATAGCCTGTGGATGGTGGTGTTCAA 600
         |||||| |||||||||||||||| ||||||||||||||| ||||||||||||||||||
Db   629 GATGACAGATATATCTGTGACCGCCTTCTACCCATGACTTGTGGTGTGGTTGTTCCAG 688

Qy   601 TTCCAGCATATAATGTGGGTCATCCTGCCCGGCATCGTCATCCTCCTGCTGTTACTGC 660
         ||||||||||||||||||||| || ||| || ||||||||||||||||||||||||||
Db   689 TTTCAGCACATCATGGTTGGCCTTATCCTGCCTGTATTCATCCTGCTCCTGCTATTGC 748

Qy   661 ATCATCATCCTTAAGTGTCACACTCCAGACAGCCTCAAGGCGCCAAGGCCCTCAAGACG 720
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   749 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAAGAGCCCAAGGCCCTCAAGACC 808

Qy   721 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGGCTGCCCATATTATGTGGGATCAGC 780
         ||||||||||||||||||||||| |||||||| |||||| ||||||||||||||||||||
Db   809 ACAGTCATCCTCATCCTGGCTTTCTTCTTTGCCTGCTGGCTGCCTTACTACATTGGGATCAGC 868

Qy   781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGGATGTGACTTCGAGAGCATTGTG 840
         ||||||||||||||||||| |||||| |||||||||||||| |||||||||||||
Db   869 ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGTGTGAGTTGAGAACACTGTG 928

Qy   841 CACAAGTGGATCTCCATCACAGAGGCCCTGCCTTCTTCCACTGTGCCTGAACCCCATC 900
         |||||||||||||| |||||||||||||||||||| |||||||||||| ||||||||||
Db   929 CACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTCTTCCACTGTTGCTGCTAAACCCCATC 988

Qy   901 CTCTATGCCTTCCTCCGGGGCCAAGTTCAAAAGCTCGCCCAGATGACACTCAACTCCATG 960
         || ||||||||| ||||| ||||||||||| ||| ||||||||||||||||| |||||||
Db   989 CTCTATGCTTTCCTTGGAGCCAAATTTAAAAACCTGCCCAGCACGCACTCCTCTGTG 1048

Qy   961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGGCGGGTGGACACTCTTCCGTC 1020
         |||||||| |||||||||||||||| |||||||||||||| |||||||||||| ||||||
Db 1049 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTT 1108

Qy 1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
         || | |||||||||||| || |||||| ||||| |||||||||||||| |
Db 1109 TCCACTGAGTCGAGTCTTCAAGTTTTCACTCCAGCTAA-------CACAGATGTAAAA 1160

Qy 1081 AATATATATATAATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG 1140
         ||||||||||||||| | ||||||||| |||| |||||| |||||| ||||||
Db 1161 GACTTTTTTTTTATACCATAAAATAAACTTTTTTTTTAAGTTACACATTTCAGATATAAAAG 1220

Qy 1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTAATTGACTGGTTGGGAGTTTATGTTCC 1200
         | ||||||||| | |||||||||||||||| |||||| ||||||| ||| |
Db 1221 ACTGACCAAATTGTACAGTTTTTTATTGCTTGTTGGATTTTTGT------CTTGTGTTTC 1274

Qy 1201 TCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAATATTGTTTTTTGTTTGTTCATGT 1260
         | ||||||||||||||| ||| |||||||||||||||||| ||||||||||||||||||
Db 1275 TTTAGTTTTTGTGAGGTTTTAATTGACTTATTATTTATA---TAAATTTTTTTTTGTTTTCATAT 1330

Qy 1261 GAATGAGCGTCTAGGCAGGCAGAACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCGTGTGTAG 1320
         |||||||| ||||||||| | |||||| |||||||||||| |||||||||||||||||||||
Db 1331 TGATGTGTGCTAGGCAGCAGCCTGTGGCCAAGTTCTTCTTAGTTGCTGTGTATGTCGTGGTAG 1390

Qy 1321 GACTGTGAGACTGTAGAGGAAGAAACTGACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
         ||||| | |||||||||| |||| |||||||||||| ||||| || ||||
Db 1391 GACTGTAGAA--------AAGGGAACTGACAATCATTCCAGAGCCGTGAGTGAATCACGTAAA 1442

Qy 1381 GCTAGCCGTGATCCCCAGCTGTTGCTGCATA 1411
         ||||||| ||| |||| ||||| ||||||||
Db 1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1473

RESULT 9
ABSS3992
ID ABSS3992 standard; cDNA; 1679 BP.
XX
AC ABSS3992;
XX
DT 02-DEC-2002 (first entry)
XX
DE DNA encoding human CXC chemokine receptor 4 (CXCR4).
XX
KW CXC chemokine receptor 4; CXCR4; chemotaxis induction; SDF1a;
KW stromal cell-derived factor 1alpha; angiogenesis; atherosclerosis;

```
Qy    61  GAACCCTGCTTCGGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC  120
          ||||||||||||||||||||||   |||||||  |||||  ||||||||||||||||||
Db   164  GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCCACCATCTAC  223

Qy   121  TTCATCATCTTCTGACTGCATGGCATGGCAATGGTGATGATCCTCGGTCATGGGTTAC  180
          ||||||||||||||  ||||||| |||||||| |||||||||||||||||||||||||
Db   224  TCCATCATCTTCTTAACTGGCATTGGGGCAATGGATTGGTCATCCTCGGTCATGGGTTAC  283

Qy   181  CAGAAGAAGCTAAGGAGCAGATGACCGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  240
          ||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||
Db   284  CAGAAGAACTGAGAAGCATGACCGGACACAAGTACCAGCAGCTGCACCTGTCAGTGGCCGACCTC  343

Qy   241  CTCTTGTCATCACACTCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG  300
          ||||||||||||||||  ||||||||||||||||||||  |||||||||||||||||
Db   344  CTCTTTGTCATCATCACGCTTCCCTTCTGGGCAGTTGATGCGTGGCAAACTGGTACTTTGGG  403

Qy   301  AAATTTTGTGTAAGGCTGTCCATATCATCTACAGCGTGCTTCTC  360
          ||||||  |||||||| ||||||  |||||  ||||||||||||
Db   404  AACTTCCTATGCAAGGCAGTCCATGTCATCTACACAGTGCAGTGTCCTC  463

Qy   361  ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAA  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   464  ATCCTGGGCCTTCATCAGTCTGGACCGGCTACCTGCCCATCGTCCACGCCACCAACAGTCAG  523

Qy   421  AGGCCAAGGAAAACTGCTGGCGCTGAAAAGGCAGTCTATGTGGGGCCTCTGGATCCCAGCCCTC  480
          ||||||||||||||||||||||||||| ||||||||||  |||||||||||||||||||||
Db   524  AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGTCTATGTTGGGCGTGATGGTGGTTGTTTCAA  583

Qy   481  CTCCTGACTATACCTGACTTCATCTTTGCGACGTCAGCCAGGGGGACATCAGTCAGGGGG  540
          |||||||||||  ||||||||||||||||||||||
Db   584  CTGGTGACTATTCCCGACTTCATCTTTGCCACG---------------TCAGTGAGGCA  628

Qy   541  GATGACAGTACATCTGTGACCGGCCTTTACCCTGACCTGGATGGTGGTGTTTCAA  600
          ||||||||||| |||||||||| ||||||||| |||||| ||||| |||||||||
Db   629  GATGACAGATATCGTGACCGGCCGTTCACCCCAATGACTTGTGGGGTTGTTGTTTCCAG  688

Qy   601  TTCCAGCATATAATGGTGGGTCCATCCTCTGCCCCGGACATCGTCATCCTCCGTGTACTGC  660
          |  ||||||||  ||||||  ||||  |||||||  |||  ||||||| ||||  ||||||
Db   689  TTTCAGCACATCATGGGTTGGCCTTATCATCCTGCTCGGTCGCTGCTATTGC  748

Qy   661  ATCATCATTCTCTAAGCTGTCACACTCCCAAGGGCCACAGAAGCGCAAGGCCCTCAAGACG  720
          || ||||||| ||||| |||||||||  ||||||||||||||||||||||||||||||||
Db   749  ATTATCATCTCCAAGCTGTCACACTCCCAAGGGCCAGGAGGCGCAAGGCCCTCAAGACC  808

Qy   721  ACAGTCAGTCCATCCTCATCCAGCTCTTCTTGCCTGCTGCGCATATTATGTGGGGGATCAGC  780
          ||||||||||||| |||  ||||||||||||||||||||  |||||||||||||||||||
Db   809  ACAGTCAGTCCATCCTCTACTCGCTTGGCTGCCTTACTACATTGGGATCAGC  868

Qy   781  ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGACATTGTG  840
          ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db   869  ATCGACTCCTTCATCCTTCATCCTCTGGGAAATCATCAAGCAAGGGGTGTGAGTTTGAGAACACTGTG  928

Qy   841  CACAAGGTCGGATCCCATCCACACAGAGGCCCTCGACCTTCTTCCACTGTTGCCTGAACCCCATC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db   929  CACAAGTGGATTTCCATCACCGGAGCCCTAGCTTGTTCTTCCACTGTTGCCTGAACCCCATC  988

Qy   901  CTCTAGCCTTCCTCCTGCGGGGGCCAAGTTCAAAAGCTCTGCCCAGCAATGCCACTCAACTCCATG  960
          |||| ||||| |||||||||||||||| |||||||  ||||||||||||||  ||||| |||  |
Db   989  CCTATGTGCTTTCCTTGGAGGCCAAAATTAAAACCTCTGCCCAGCACCTCTGCTGTG  1048

Qy   961  AGCAGAGAGGTCCAGCGCTCCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  1049  AGCAGAGAGGTCCAGCCTCCAAGATCCTCTCCCAAAGGAAAGCGAAGAGGAGGGACATTCATCCTGTT  1108

Qy  1021  TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTCTAACCCTTATGCAAAGACTTATAT  1080
          ||||||  ||||||||||||||||| ||||||  |||||||||||||||||||||||||
Db  1109  TCCACTGTCTGAGTCTTCAAGTTTCACTCCAGCTAA------CACACAGTGTAAAA  1160

Qy  1081  AATATATATATATATGATAAAGAACTTTTTTTATGTTACACATTTCCAGATATAAGAG  1140
          ||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db  1161  GACTTTTTTTATACGATAAATAACTTTTTTTTAAGTTACACATTTTCAGATATAAAAG  1220

Qy  1141  ACTGACCAAGTCTTGTACAGTTTGTTGTTTTTTTTTTTTTTAATTGACTGTGGGAGTTTATGTTCC  1200
```

```
Db  1221  ACTGACCAATATTGTACAGTTTTTATTGCTGTGGATTTTGT------CTGTGTTTC  1274
          ||||||||| |||||||||||| |||||  ||||||||||||      ||| ||||
Qy  1201  TCTAGTTTTGTGAGGTTTGACTTAATTATTATAAATATTGTTTTGTTTTCATGT  1260
          |||||||||||||||||||||| |||| |||||||---|||||||||||||||||
Db  1275  TTTAGTTTTGTGAGTTAAATGACTTATTATA----TAAATTTTTGTTTTCATAT  1330

Qy  1261  GAATGAGGCTCAGGCAGGAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTATCGTGTGTAG  1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1331  TGATGTGTGTTAGGCAGGACCACCTGTGGCCAAGTTCTTAGTTGCTGTATGTCTCCGTGTGTAG  1390

Qy  1321  GACTGTAGAACTGTAGAGGAAGAAGAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAA  1380
          ||||||||||       ||||||||||||||||||||||||||||||||||||||
Db  1391  GACTGTAGAA---------AAGGGAACTGAACATTCCAGAGCGTGTAGTGAATCACGTAAA  1442

Qy  1381  GCTAGCCGTGATCCTCAGCTGTTGCTCGCATA  1411
          ||||||||||||||||||||||| ||||||||
Db  1443  GCTAGAAATGATCCCCCAGCTGTTTATCCATA  1473
```

```
RESULT 10
ABV78039
ID   ABV78039 standard; DNA; 1679 BP.
XX
AC   ABV78039;
XX
DT   12-NOV-2002  (first entry)
XX
DE   Hypoxia-regulated protein coding sequence #59.
XX
KW   Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
KW   antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;
KW   hypoxia; tumourigenesis; angiogenesis; apoptosis; cancer;
KW   ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
KW   preeclapmsia; atherosclerosis; inflammatory condition; wound healing;
KW   inflammation; erythropoiesis; hair loss; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO20024665-A2.
XX
PD   13-JUN-2002.
XX
DR   10-DEC-2001; 2001WO-GB005458.
XX
PF   08-DEC-2000; 2000GB-00030076.
PR   08-FEB-2001; 2001GB-00003156.
PR   25-OCT-2001; 2001GB-00025666.
XX
PA   (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI   White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
PI   Rayner WN;
XX
DR   WPI; 2002-627238/67.
XX
PT   Identifying a gene involved in disease for treating hypoxia-regulated
PT   conditions, comprises comparing the transcriptome/proteome of two cell
PT   types under different conditions and identifying a differentially
PT   regulated gene.
XX
PS   Claim 37; Page 378; 538pp; English.
XX
CC   The present invention relates to methods for identifying genes and
CC   proteins that are implicated in a specific disease or physiological
CC   condition. The method comprises comparing the transcriptome/proteome of a
CC   specialised cell type implicated in a disease or condition with that of a
CC   second specialised cell type, under two experimental conditions, and
CC   identifying a gene that is differentially regulated in the two
CC   specialised cell types under experimental conditions. ABV77873-ABV78116
CC   and ABP65061-ABP65257 were identified using the methods of the invention.
CC   The coding sequences and proteins are useful for treating a disease in a
CC   patient, for manufacture of a medicament for treating hypoxia-regulated
```

```
CC    conditions, and for regulating tumourigenesis, angiogenesis, apoptosis,
CC    biological response to hypoxia conditions, or hypoxic-associated
CC    pathology in a patient. The coding sequences and proteins are also useful
CC    for monitoring the therapeutic treatment of a disease or physiological
CC    condition, such as cancer, ischaemic conditions, reperfusion injury,
CC    retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory
CC    conditions, wound healing, inflammation, erythropoiesis or hair loss
XX
SQ    Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match             55.1%;  Score 933.4;  DB 6;  Length 1679;
Best Local Similarity   81.8%;  Pred. No. 4.3e-207;
Matches 1154;  Conservative    0;  Mismatches  216;  Indels   41;  Gaps    5;

Qy      1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
            ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    104 ATATACACTTCAGATAACTACACACCGAGGAAATGGGCTCAGGGACTATGACTCCATGAAG 163

Qy     61 GAACCCGCTTCCGGGATGAAAAGCTTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    164 GAACCCGCTTCCGTGAAGAAAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTAC 223

Qy    121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGTCATGGGTTAC 180
            |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    224 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTGATCATCCTGTCATGGGTTAC 283

Qy    181 CAGAGAAGTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    284 CAGAGAAACTGAGAAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC 343

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGGTGACTTTTGGG 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 CTCTTTGTCATCACGCTCCCCTTCTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTTGGG 403

Qy    301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 AACTTCCTATGCAAGGCAGTCATCATCATCTACACTGTCAACCTTTACACAGCAGTGTCCTC 463

Qy    361 ATCCGGCTCATCAGCAGCCTGGACCGGACCGTGACTTCGCCATTGTCACGCACCAGTCAA 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    464 ATCCTGGCCTTCATCAGTCGCACCAGTCTGGACCCATCCGGCCATCCTGGCCATCCCAGCCCTC 523

Qy    421 AGGCCAAGGAAACTCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    524 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTCATGTTGGGCGTCTGGATCCCTGCCCTC 583

Qy    481 CTCCTGACTATACCTTGACTTCATCTTGCCGACGTCAGCGAGGGAACATCAGTCAGGGG 540
            |||||||||||||||||||||||||||||||||||||||||||||        ||||||
Db    584 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG------------TCAGTGAGGCA 628

Qy    541 GATGACAGGTACATCTGACGCCGCCTTACCCGATAGCCTGTGGATGTGGTGTTTCAA 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    629 GATGACAGATATATCTGTTGGACCGGTTGCACCCCAATGACTTGTGGGTGGTGTTCCAG 688

Qy    601 TTCCAGCAGATAATAATGGTGGGTCTCATCCTGCCGGCATCGTCATCCTCCCTGTTACTGC 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    689 TTTCAGCAACATCATCGGCGCCTTATCCTGCCTTGCTATTGTCATCCTGTCCTGCTATTGC 748

Qy    661 ATCATCATCCTAGCTGTCACACTCCAAGGGCCACCAGGAAGCGGCAAGGCCCTCAAGACG 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    749 ATTATCATCTCCCAAGCTGTCACACTCCAAGGGCCACCAGGAAGCGGCAAGGCCCTCAAGACC 808

Qy    721 ACAGTGCATCCTCATCCTAGCTTCTTTGCCTGCTGGCTGCTCCATATTATGTGGGATCAGC 780
            |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    809 ACAGTGCATCCTCATCCTCCTCACCTGGCTTCTGGCTGCTCGGCTGTTGGCTGCTCCATACATTGGGATCAGC 868

Qy    781 ATCGACTCCTTCATCCTTTGGGAACTCATCAAGCAAGGATGTGACTTCGAGAGGCATTGTG 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    869 ATCGACTCCTTCATCCTTCCTGGAAATCATCAAGCAAGGAGGTGTGAGTTTGAGAACACTGTG 928

Qy    841 CACAAGTGGATCTCCATCACCAACAACGAGGCCCTCGCCTTCTTCTTGCCTGCTGCTCAGCCCCATC 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    929 CACAAGTGGATTTCCATCACCAACAACGAGGCCTAGCTTCCTTCTTGCTGTTCTGACCCCATC 988

Qy    901 CTCTTATGCCTTCCTCCTCGGGGCCCAAGTCAGCTCGGGGCCCAGCATGGCATGCGACTCAACTCCATG 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    989 CTCTATGTTTCCTTGAGCCCAAAATTTAAAACCTCTGCCCAGCACCGGACTCACCCTCTG 1048

Qy    961 AGCAGAGGCTCCAGCTTCAAGATCCTTTCCAAAGAGAAAGCGGGGTGGACACTCTTCCGTC 1020
            |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   1049 AGCAGAGGGTCCAGCTTCAAGATCCTTCCAAAGGAAGCGGAGGCGGGACAATTCATCTGTT 1108

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCGAAAGACTTATAT 1080
            |||||||||||||||||||||||||||||||||||||||||||||      ||||||||||
Db   1109 TCCACTGAGTCTGAGTCTTCAAGTTTTCACTCCCAGCTAA------CACAGATGTAAAA 1160

Qy   1081 AATATATATATATATTATATATGATAAAGAAACTTTTTATGTTACACATTTTCAGATATAAGAG 1140
            || || |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   1161 GACTTTTTTTATACGATAAATAACTTTTTTTTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1220

Qy   1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
            ||||||||||||||||||||| |||||||||     ||||||||
Db   1221 ACTGACCAATATTGTACAGTTTTTTAATTCTTGTGTGGATTTTGT-----CTTGTGTTTC 1274

Qy   1201 TCTAGTTTTTGTGGAGGTTGACTAATTTATATAAATATTGTTTTTTTTGTTTTTCATGT 1260
            ||||| |||||| |||||||||||||||||      ||||||||||||||||||||||
Db   1275 TTTAGTTTTTGTGGAAGTTAATTAATTGACTTATTTATA----TAAATTTTTTTTGTTTCATAT 1330

Qy   1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1331 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTCCTTAGTTGCTGTATGTCTCGTGGGTAG 1390

Qy   1321 GACTGTAGAACTGTAGAGGAGAAGAAAACTGAACATTCCAGAATGTGTGGGTAAATTGAATAAA 1380
            |||||||||          |||||||||||||||||||||||||||||
Db   1391 GACTGTAGAA--------AAGGGAACTGAACATTCCAGAGCCTGTTTATGCATA 1442

Qy   1381 GCTAGCCCTGATCCTCAGCTGTTGCTGCTCATA 1411
            |||||||||||||||||||||||||||||||||
Db   1443 GCTAGAAATGATCCCCAGCCTGTTTATGCATA 1473
```

```
RESULT 11
ABX08779
ID   ABX08779 standard; cDNA; 1679 BP.
XX
AC   ABX08779;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Angiogenesis-associated human polynucleotide sequence #41.
XX
KW   Human; angiogenesis-associated transcript; angiogenesis;
KW   angiogenesis-associated disease; cancer; cytostatic; gene therapy; gene;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200279492-A2.
XX
PD   10-OCT-2002.
XX
PF   14-FEB-2002;  2002WO-US004915.
XX
PR   14-FEB-2001;  2001US-00784356.
PR   22-FEB-2001;  2001US-00791390.
PR   19-APR-2001;  2001US-0285475P.
PR   03-AUG-2001;  2001US-0310025P.
PR   13-NOV-2001;  2001US-0350666P.
PR   29-NOV-2001;  2001US-034244P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Murray R, Glynne R, Watson SR, Aziz N;
XX
DR   WPI; 2003-040681/03.
DR   P-PSDB; ABU03495.
```

```
XX
PT   Detecting angiogenesis-associated transcript in a cell for diagnosing and
PT   treating cancer by contacting a sample with a polynucleotide that
PT   exhibits changes in expression level as a function of time in tissue
PT   undergoing angiogenesis.
XX
PS   Example 2; Page 217-218; 291pp; English.
XX
CC   The present invention relates to methods and compositions for detecting
CC   an angiogenesis-associated transcript in a cell in a patient. The method
CC   involves contacting a biological sample from the patient with a
CC   polynucleotide that selectively hybridises to a sequence at least 80%
CC   identical to any of the angiogenesis-associated human polynucleotide
CC   sequences given in the specification. These angiogenesis-associated
CC   polynucleotide sequences comprise genes that exhibit changes in
CC   expression levels as a function of time in tissue undergoing
CC   angiogenesis. The method and the polynucleotide sequences of the
CC   invention are useful for diagnosing and treating angiogenesis and
CC   angiogenesis-associated diseases e.g. cancer. The polynucleotide
CC   sequences are also useful in the gene therapy of such disorders. The
CC   angiogenesis-associated proteins encoded by the polynucleotide sequences
CC   are useful as a vaccine for therapeutic and prophylactic immunisation.
CC   ABX08739-ABX08853 represent angiogenesis-associated polynucleotide
CC   sequences
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match    55.1%;   Score 933.4;   DB 7;   Length 1679;
Best Local Similarity   81.8%;   Pred. No. 4.3e-207;
Matches 1154;   Conservative   0;   Mismatches 216;   Indels   41;   Gaps   5;

Qy         1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
             |||||||  ||||| ||||||||||||||  |||| |||||||||||||||||||||||||
Db       104 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCCAGGGGACTATGACTCCATGAAG 163

Qy        61 GAACCCTGCTTCCCGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
             |||||||||||||  |||||||  ||||| |||||||||||||  ||||||||||||||||
Db       164 GAACCCTGTTCCCGGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC 223

Qy       121 TTCATCATTCTTGACTGACTAGTCGGCAATGGTGATTGGTCATCCTGTCATGGGTTAC 180
             |||||||||||| |||||||  ||||||||||||||  ||||||||||||||||||||||
Db       224 TCCATCATTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGTCATGGGTTAC 283

Qy       181 CAGAGAGAGCTAAGGAGCATGACGGACAAGTACCGGCGTCACCTGTCAGTGGCTGACCTC 240
             |||||||||||||||||||||||  ||||||  ||||||||||||||||||||||||||||
Db       284 CAGAGAAGAACTGAGAAGCAGAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC 343

Qy       241 CTCTTTGTCATCACACTCCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 300
             ||||||||||||| || |||||||| |||||| ||  ||||||||||||||||||||||||
Db       344 CTCTTTGTCATCACGTCCTTCCCTTCTGGGCAGTTGATGCCGTGGCTGGCAAACTGGTACTTTTGGG 403

Qy       301 AAATTTTGTGTAAGGCTGTCCATATCATCCATCTACACTGTCCACTGCTACAGCAGCGTTCTC 360
             ||||  |||||  ||||||  |||||||||||||| |||||||||||||||||||||||||||
Db       404 AACTTCCTATGCAAGGCAGTCTGGCTGGTTGGCTGAAAAGGTGGTCTATGCTCTTGGGCGTGGCTTCCTC 463

Qy       361 ATCCTGGCCTTCATCAGCCTGGACTGGACCTACCTTGCCATTGTCCACGCCACCAACAGTCAA 420
             ||||||||||||||||||||||  |||||  ||||| |||||||||||||||  |||||||||
Db       464 ATCCTGGCCTTCATCAGCTGACGTCCGACCCGCTTCATCGTCCATCGTCACGCCACCAACAGTCAG 523

Qy       421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
             |||||| || ||||||||  ||||||||||  ||||||||||  ||||||||||||||||
Db       524 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTTGGCGTCTTGGATCCCTGCCCTC 583

Qy       481 CTCCTGACTATACCTGACCTCATCTTGCCGACGTCAGCCAGCCAGGGGACATCAGTCAGGGGG 540
             |||||||||||||||||||||||||||||||||          |||||  |||||||||||
Db       584 CTGCTGACTATTCCCGACTTCATCTTTGCCACCG--------TCAGTGAGGCA 628

Qy       541 GATGACAGAGTACATCTGTCGACCCGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA 600
             ||||||||  ||||||||  || ||||||||||||||||||||  |||||||||||||||||
Db       629 GATGACAGATATATCTGTGGACCCGTTCCACCCCAATGACTTGTGTGGGGTGGTTGGTGTTCCAG 688

Qy       601 TTCCAGCATATAATGGTGGGTCTCATCCTCCCGGCATCGTCATCCTCCTGTTACTGC 660
             || ||||| ||  |||||||   |||||||  ||||| ||||||||  || ||| ||
```

```
Db       689 TTTCAGACAATCATGGTTGGCCTTATCTGCCTTATCTGCCTGGTATTGTCATCCTGCTGCCGTATTGC 748

Qy       661 ATCAATCATCTCAAGTGCACACTCCAGAATGCCACCAGAGAGGCCAAGGCGCAAGAGCGCAAGGACACG 720
             ||| ||||||||||| ||  |||||| |||||||||||||| |||||||||||||||||||||||||||
Db       749 ATTATCATCTCCAAGTGTCACACTCCAGAGGGCCACCAGAGGGCCACCAGAGAGGCCAAGGCGCAAGACC 808

Qy       721 ACAGTCATCCTCATCCTAGCTTCTTTGCCTGCGTCCCATATATATGTGGGATCAGC 780
             |||||||||||| ||||||||||||||||||||||||  |||||||||||||||||
Db       809 ACAGTCATCCTCATCCTGGCTTCTTTCTGCGTGGCTGCCTACTACATTACACAATTGGGATCAGC 868

Qy       781 ATCGACTCCTTCATCTTTTGGGAGTCATCAAGCAGGATGTGACTTCGAGAGCATTGTG 840
             ||| ||||||  |||| |||||||| |||||||||||||  |  |||| ||||||||||
Db       869 ATCGACTCCTTCATCTTTTCATCCTGGAAATCATCAGCAAGGGTGTGAGTTTGAGAACACTGTG 928

Qy       841 CACAAGTGGATCCATCACAGAGAGGCCCTCGGCCTTCTTCCACTGTTGCCTGACCCCATC 900
             |||||||||||| ||| |||||||||| ||  ||||| ||  |  ||||||| ||||||||
Db       929 CACAAGTGGATTTTCATCACCAGGAGCCCTAGCTTTCCTTCCCACTGTGTCTGGTAACCCCCATC 988

Qy       901 CTCTATGCCTTCCTGGGGCCCAAGTCAAAGCTCTGCCCAGCCATGCCACTCAACTCCATG 960
             |||||||||| | ||||||||||||||||  ||||  ||||||||||||||| ||||||||
Db       989 CTCTATGCCTTTCCTTGAGCCCAAATTTAAAACCTCTGCCCCAGCCACCACCACCTGTG 1048

Qy       961 AGCAGAGGCTCCAGCCTCAAGATCCTTCAAAGGAAAGCCGGGTGGGTCGGACACTCTTCCGTC 1020
             ||||||||| ||||||||| ||||  ||||||| ||||||||||| ||||||||||||  |||
Db      1049 AGCGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGGCAGGTCGGAGGTGGACATTCATCCTGTT 1108

Qy      1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCAGCTGACCCTTATGCAAAGACTTATAT 1080
             ||||  ||||||||||||||||||| ||||| ||||| ||||||      |||||||||
Db      1109 TCCACTCGAGTCGAGTCTTCAAGTTTCACTCCAGCTAA--------CACAGATGTAAAA 1160

Qy      1081 AATATATATATATGATAAAGACTTTTTTATGTTACACATTTTCCAGATATAAGAG 1140
             || |||  |||||  |||||||||||||| ||| || |||||||| | |||||||
Db      1161 GACTTTTTTTATGAAATAAATAACTTTTTTTTTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1220

Qy      1141 ACTGACCAGTCGTGTACAGTTTTTTGTTGTTGGGAGTTTATGTTCC 1200
             ||||||||  |||||| |||||  |||||||          |||||||
Db      1221 ACTGACCAATATTGTACAGTTTTTTATTGCTTGGTTGTTGTTGT-----CTTGTGTTTC 1274

Qy      1201 TCTAGTTTTTGTGAGGTTGACTAATTATTATAAATATTGTTTTTTGTTTTTGTTTCATGT 1260
             ||| ||||||||||||||  |||||||||| |      |||| ||||||||||| |||||
Db      1275 TTTAGTTTTTTGTGAAGTTGTTAATTGACTTATTTATA---TAAATTTTTTTTTGTTTTTCATAT 1330

Qy      1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
             |||||||||  |||| |||  ||||||  |||||||| |||||||||||||||||||||||
Db      1331 TGATGTGTGTCTAGGCAGGCAGACCTGTGGGCCAAGTTCTTAGTTGCTGTTTATGTCTCGTGGTAG 1390

Qy      1321 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAA 1380
             |||||||||                 |||||| ||||||||||||||||||||||||||
Db      1391 GACTGTAGAA--------AAGGGAACTGAACATTCCCAGAGTGTTTTATGCATA 1442

Qy      1381 GCTAGCCGTGATCCTCCAGCTGTTGCCTGCATA 1411
             ||||||  |||||||| |||| |||| |||||||
Db      1443 GCTAGAAATGATCCCCCAGCTGTTTATGCATA 1473
```

```
RESULT 12
ABX74454
ID   ABX74454 standard; cDNA; 1679 BP.
XX
AC   ABX74454;
XX
DT   21-MAR-2003   (first entry)
XX
DE   Human cDNA sequence #26 up-regulated in CC-RCC patients.
XX
KW   Human; microarray; solid surface; immobilised probe; CC-RCC;
KW   differential expression profile; aggressive CC-RCC tumour type;
KW   non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
KW   gene expression profiling; tumour tissue; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO20029411-A2.
```

XX
PD   10-OCT-2002.
XX
PF   29-MAR-2002; 2002WO-US009576.
XX
PR   29-MAR-2001; 2001US-0279411P.
XX
PA   (VAND-) VAN ANDEL INST.
XX
PI   Haab B, Rhodes D, Teh BT, Takashi M;
XX
DR   WPI; 2003-040679/03.
XX
PT   New microarray, comprising a matrix of cDNA probe from a set of probes
PT   immobilised to a solid surface in predetermined order, useful in the
PT   prognosis of patients with clear cell renal carcinoma.
XX
PS   Claim 35; Page 132-133; 179pp; English.
XX
CC   The present invention relates to a microarray comprising a matrix of at
CC   least one cDNA probe from a set of probes immobilised to a solid surface
CC   in a predetermined order, where a row of pixels corresponds to replicates
CC   of one distinct probe from the set. The probes are complementary to
CC   nucleic acid sequences that are expressed differentially in aggressive as
CC   compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
CC   and that hybridise to the probes under high stringency conditions. The
CC   microarray is useful for the prognosis of patients with CC-RCC, wherein
CC   aggressive and non-aggressive CC-RCC tumour types are characterised by
CC   differential expression profiles of genes that hybridise with one or more
CC   probes immobilised on the microarray. The arrays are useful for gene
CC   expression profiling of tumour and normal tissues. The present sequence
CC   represents a human cDNA sequence up-regulated in CC-RCC patients
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match   55.1%; Score 933.4; DB 7; Length 1679;
Best Local Similarity   81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

```
Qy      1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGCTCTGGAGACTATGACTCCAACAAG  60
          ||||||||||||| ||||||||||||||||||||||||| || ||||||||||||||
Db    104 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163

Qy     61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 120
          |||||||| ||||||||||||||| |||||||||||| ||||||||||||||||||
Db    164 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTAC 223

Qy    121 TTCATCATCTTCTTGACTGGCATAGTCGCAATGATTGGTGATCCTGGTCATGGGTTAC 180
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    224 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGATTGGTCATCCTGGTCATGGGTTAC 283

Qy    181 CAGAAGAAGCTAAGGAGCATGAACGGACAAGTACCCGGCTGACGCTGTCAGTGGGCTGACCTC 240
          |||||||||||||| |||||||| |||||||| |||||||||||||||||||||||
Db    284 CAGAGAAACTGAGAAGCATGACGGACACAGGCTGCACCTGTCAGTGGGCCGACCTC 343

Qy    241 CTCTTTGTCATCACACTCCCCTTCGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 300
          |||||||||||||||||||||||||||||||| |||||||| |||||||||||||||
Db    344 CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCGTGGCTGACTGGTACTTTGGG 403

Qy    301 AAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCAGCGTTCTC 360
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    404 AACTTCCTATGCAAGGCAGTCCATGTCATCTACACAGTCAACCTCTACAGCAGTGTCCCTC 463

Qy    361 ATCCTGGCCTTCATCAGCTGGACCGGTACCTCGCCATTGTCCACGCCACCAGTCAA 420
          |||||||||||||| ||||||||||||||||||||||||| ||||||||||||||
Db    464 ATCCTGGCCTTCATCAGTCTGGACCGCTGGACCCGTGGGCCGTCTGAATGTGGGCCCTC 480

Qy    421 AGGCCAAGGAAAACTGCTGGTGGGCTGAAAAGGCAGTCTATGTGGCCGTCTCAGCCCCTC 480
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    524 AGGCCAAGGAAGCTGTGTGGGCTGAAAAGGTGTCTATGTTGGCCGTCTGGAGTCCCCC 583

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACAATCAGTCAGGGG 540
          ||||||||||||||||| |||||| ||
```

```
Db    584 CTGCTGACTACTATTCCGACTTCATCTTTGCCACG----------TCAGTGAGGCA  628
          |||||||||||||||||||||||||||||| ||||
Qy    541 GATGACAGTACATCTGTGACCGGCCTTTACCCGATAGCCTGTGGGATGGTGGTGTTCAA  600
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||| |||||
Db    629 GATGACAGATATATCTGTGACCGGCCTTTCACCCCAATGACTTGTGGGTGTGTTGTTCCAG  688

Qy    601 TTCCAGCATATATGGTGCGTCTCATCCTGCCGGTCATCGTCATCCTCTCCTGTTACTGC  660
          ||||| |||| |||||| |||||||||||||| ||||||||||||||||||| |||||
Db    689 TTTCAGCACCATCATGGTTGGCCTTATCCTGCCTGGTATTGTCATCCTGTCCTGCTATTGC  748

Qy    661 ATCATCATCTCTAAGCTGTCACACTGCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACG  720
          ||||||||||||||||| ||||||||||||||||||||||| |||||| |||||||||
Db    749 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACC  808

Qy    721 ACAGTCATCATCCTGGCTTTCTTGCTCTGCTGGCTGCCATATATGTGGGGATCAGC  780
          |||||||||||||||||| ||||| |||||||||||||||||| |||||||||||||
Db    809 ACAGTCATCCTCATCCTGGCTTTCTTCGCCTGTTGGCTGCCTTACACCATTGGGATCAGC  868

Qy    781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  840
          |||||||||||| |||||| |||||||||||||||||||||||||||| ||| |||||
Db    869 ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG  928

Qy    841 CACAAGTGGATTCCATCACAGAGGCCCTCTTCCACTGTGCCTGAACCCATC  900
          ||||||||| |||||||| ||||||||||| ||||||| |||||||||||||
Db    929 CACAAGTGGATTTCCATCACCGAGGCCCTAGCTTCTCCACTGTGCTGAACCCCATC  988

Qy    901 CTCTATGCCTTCCTCGGGGCCAGTTCAAAAGTCTGCCAGCATGCACTCAACTCCATG  960
          ||||||| |||||||||||| |||||||||||||| |||||| ||| |||||||||
Db    989 CTCTATGCTTTCCTTGGGAGCCAAATTTAAAACCTCTGCCCAGCACGCACCACCTCTGTG 1048

Qy    961 AGCAGAGGGCCTCCAGCCTCAAGATCCTCCTCCAAAGGAAAGCGAGGTGGACACTTGGACATTCATCTGTT 1020
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   1049 AGCAGAGGGTCCAGCCTCAAGATCCTCCTCCAAAGGAAAGGAGGGCAGGTGGACACTCGTT 1108

Qy   1021 TCCACGGAGTCGAGAATCCTCCAGTTTCACTCCCAGTCAAACCCTTATGCAAAGACTTATAT 1080
          |||||||||||||||||||| |||||||||||||||||  ||
Db   1109 TCCACTGAGTCTGAGTTTCAGTTTTCACTCCCAGTAA---------CACAGATGTAAAA 1160

Qy   1081 AATATATATATATATATATGATAAAGAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG 1140
          ||||||||||| |||||| ||||||| |||||||||||||||||||| |||||||||
Db   1161 GACTTTTTTTTATACGATAAATAAAACTTTTTTTTTTTTTTAAGTTACAACATTTTCAGATATAAAAG 1220

Qy   1141 ACTACCAGTCTTGTACAGTTTTTTTTTTTTTTTGACTGTTGGGGAGTTTATGTTCC 1200
          ||||||||||||||||||||||||        |||||| |||
Db   1221 ACTGACCAATATTGTACAGTTTTTATTGCTTGTGTGGATTTTTGT------CTTGTGTTTC 1274

Qy   1201 TCTAGTTTTTGTGGAGGTTGACTTAATTATATAAAATATTGTTTTTTGTTTTCATGT 1260
          |||||||| |||||||||||| ||| |||||||||||| |
Db   1275 TTTAGTTTTTGTTGAAGTTTAATTGACTATTTATA----TAAATTTTTTTGTTCATAT 1330

Qy   1261 GAATGAGCGTCAGGGAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
          |||||||| ||||||||||||||||| |||||||||||||||| ||||||||||||
Db   1331 TGATGTGTGTCTAGGCAGGACCTGTGCTGGTCTTCAGTTCTTAGTTGCTGTATGTCTCGTGGTAG 1390

Qy   1321 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAA 1380
          |||||||||      |||||||||||||||||||| |||||||||||||||
Db   1391 GACTGTAGAA--------AAGGGAACTGAACATTCCAGACTGTTATGCATA 1442

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
          ||| |||||||| ||||||||||||||
Db   1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1473
```

RESULT 13
AB268886
ID   AB268886 standard; cDNA; 1679 BP.
XX
AC   AB268886;
XX
DT   28-MAY-2003 (first entry)
XX
DE   Nucleotide sequence of human chemokine receptor CXCR4.
XX

```
KW   Human; chemokine receptor; CXCR4; viral infection; surface protein;
KW   respiratory virus infection; respiratory syncytial virus infection;
KW   RSV infection; bronchiolitis; bronchitis; pneumonia; asthma; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             89..1147
FT                   /*tag= a
FT                   /product= "CXCR4"
XX
PN   WO200304153-A2.
XX
PD   20-FEB-2003.
XX
PF   12-AUG-2002; 2002WO-CA001248.
XX
PR   10-AUG-2001; 2001US-0311088P.
XX
PA   (TOPI-) TOPIGEN PHARM INC.
XX
PI   Renzi P, Zemzoumi K;
XX
DR   WPI; 2003-256541/25.
DR   P-PSDB; ABP97733.
XX
PT   Modulating viral infection of a cell, for treating or preventing
PT   respiratory virus infections, bronchitis, pneumonia or asthma, by
PT   modulating a binding interaction between a cell chemokine-receptor and a
PT   surface protein of the virus.
XX
PS   Disclosure; Page 117-119; 120pp; English.
XX
CC   The present sequence encodes human chemokine receptor CXCR4. The
CC   specification describes a method for modulating viral infection of a
CC   cell. the method comprises modulating a binding interaction between a
CC   cell chemokine-receptor and a surface protein of the virus. The proviso
CC   is that the cell chemokine-receptor is not CX3CR1 and that the virus is
CC   not HIV. The method is useful for treating or preventing respiratory
CC   virus infection in vertebrates, more particularly respiratory syncytial
CC   virus (RSV) infections, and related diseases, e.g. bronchiolitis,
CC   bronchitis, pneumonia or asthma
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
```

```
Query Match             55.1%;  Score 933.4;  DB 7;  Length 1679;
Best Local Similarity   81.8%;  Pred. No. 4.3e-207;
Matches  1154;  Conservative  0;  Mismatches  216;  Indels  41;  Gaps  5;
```

```
Qy     1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGGTCTGAGAGACTATAGACTCCAACAAG  60
         |||||||||| ||||||||| ||||||||||||||| |||
Db   104 ATATACACTTCAGATAACTAACTACACCGAGAAATGGGCTCAGGGGACTATAGACTCCATGAA 163

Qy    61 GAACCCTGCTTCCGGGATGAAGAAACGTCCATTCAATAGGAATCTTCCTGCCACCATCTAC 120
         ||||||| ||||  ||||||||||||||||||||||
Db   164 GAACCCTGTTTCCGTGAAGAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC 223

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTTCATGGGTTAC 180
         |||||||||| ||||||||||
Db   224 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 283

Qy   181 CAGAAGAAGCTAAGGAGCATGACGACGACGACAAGTACGGCTGCACCCTGCAGTGGCTGACCTC 240
         ||||||||| ||||||||||
Db   284 CAGAAGAAACTGAGAAGCATGACGACGACGACAAGTACGGCTGCACCCTGCAGTGGCTGGGCCGACCTC 343

Qy   241 CTCTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 300
         |||||| |||||||||
Db   344 CTCTTTGTCATCACGCGCTTCCCTTCTGGGCAGTTGATGCCGCAGTTGAATCCGTACTTTGGG 403

Qy   301 AAATTTTTGTGTAAGGCTGTCCATATCATCATCACACTGTCAACCCTCTACAGCAGCGTTCTC 360
         ||||| ||||||||
Db   404 AACTTCCATGCAGAGGCAGTCATGCATGTCATCATCTCAACACTCACAGTCACAGCAGTGTCCTC 463
```

```
Qy   361 ATCCTGGCTTCATCAGCTGGACCGGTACCTTCGCCATTGTCCAGCGCCACCAACAGTCAA 420
         |||||||  ||||||| ||||||||||  |||  ||||||||||| ||||||||||||||
Db   464 ATCCTGGCCTTCATCAGTCTGGACCGGCCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAG 523

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGGGTCTGGATCCCGACCCCTC 480
         |||||||||||| || ||||||||||  ||||||| ||||  |||||||||| |||||||
Db   524 AGGCCAAGGAAGCTGTTGGCTGAAAAGTGGTCTATGTTGGCCTCTGGATCCCTGACCCTC 583

Qy   481 CTCCTGACTATACCTGACTCATCTTTGCCGACGTCAGCGACATCAGCGACGGG 540
         || |||||||| || ||||||||||||||||||        |||| ||| ||
Db   584 CTGCTGACTATTCCCGACTCATCTTTGCCAACG------TCAGTGAGGCA 628

Qy   541 GATGACAGGTACATCTGTGACCGCCTTACCCGATAGCCTGTGGATGGTGGTGTGTTCAA 600
         |||||||| ||| ||||||||||||||  |||| |||||  ||||||| |||||||||
Db   629 GATGACAGATGATATCTGTGACCGCCTTCTACCCCAATGACTTGGGGTGGTTGTTGTTCCAG 688

Qy   601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTGTTACTGC 660
         || |||||  ||||| || |||  || |||||  || ||||||||||| ||||| |||
Db   689 TTTCAGCACATCATGGTTGGCCTTACTGCCTTATCCTGCCTGTATTGTCATCCTGCTATTGC 748

Qy   661 ATCATCATCTCTAAGCTGTCACACTCCAAGAGGCACCAGAAGGCAAGGCCCTCAAGACG 720
         || |||||||| |||||||||||||||||||||| |||||||| |||||||||||||||
Db   749 ATTATCATCTCCAAGCTGTCACACTCCAAGAGGCCACCAGAAGCGCCAAGGCCCTCAAGACC 808

Qy   721 ACAGTCATCCTCATCCTAGCTTTCTTGCCTGCTGGTGGCCATATATATGTGGGGATCAGC 780
         |||||||| ||||||||||||||||| ||||||| ||||| ||||| |||||||||||||
Db   809 ACAGTCATCATCCTCATCCTGGCTTTCTTCTTCTGCCTGTGGGCCTGTGCCTTACTCACATTGGGATCAGC 868

Qy   781 ATCGACTCCTTCATCCTTTGGAGTCATCAAGGATGTCGACTTCGAGAGCATTGTG 840
         |||||||||||||||||  |||| ||||||| || |||| |||||||||||||||
Db   869 ATCGACTCCTTCATCCTCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGACACTGTG 928

Qy   841 CACAAGTGGATCTCCATCACAGAGAGCGCCCTCCTTCTCCACTGTTGCCTGAACCCCATC 900
         |||||||||||| |||||||||  ||||||||| ||||||||| ||||| |||||||||
Db   929 CACAAGTGGATTTCCATCACACCGAGGCCCTCAGCTTTCTCCACCGTTGTCTGAACCCCATC 988

Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTCAAAAGCTCTGCCCAGCATGCACTCAACTCACTCATCCATG 960
         |||||||| ||||| || |||||  | |||| ||||||||||| || |||| |||| ||||||
Db   989 CTCTATGCTTTCCTTGGAGCCAAATTAAAACCTCTGCCCAGCACGCGACCACTCACCTCTGTG 1048

Qy   961 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAGTCGGGTGGCCACTCTTCCCGTC 1020
         |||||||| |||||||||||||||||  |||||||||||  |||| ||| |||  |||||
Db  1049 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAGGAAGCGAGGTGGAGGTGGACGTGGACATTCATCCTGTT 1108

Qy  1021 TCCACGGAGTCAGAATCCTCCCAGTTTCACTCCAGCTAACCCTTATGCGAAAGACTTATAT 1080
         ||||| |||| ||| ||||| ||||||||||||||||||                      
Db  1109 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAA------CACAGATGTAAAA 1160

Qy  1081 AATATAATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG 1140
         |  ||||| |||||| || |  ||||| ||||| ||||||||||| |||||||| |||
Db  1161 GACTTTTTTTTTATCAGATAAATAATAACTTTTTTTTTTTTTTAAGTTACACATTTTTCAGATATAAAAG 1220

Qy  1141 ACTGACCAGTCTTGTACAGTTTTTTTAATTGACTGTTGGGGAGTTTATGTTCC 1200
         |||||||| | ||||||||||||| || |||||||| |||| ||||| ||||
Db  1221 ACTGACCAATATGTACAGTTTTTATTGCTGTGTTGGATTTTTGT----CTTGTGTTTC 1274

Qy  1201 TCTAGTTTTTGTGTGGAGGTTTGACTTAATTTATATAAATATATTGTTTTTGTTTCATGT 1260
         | ||||||||| ||| ||||||| |||| ||||||    |||||  |||||||||||| |
Db  1275 TTTAGTTTTTTGTGAAGTTTTAATTGACTTATTTATA----TAAATTTTTTTTGTTTCATAT 1330

Qy  1261 GAATGAGCGTCTAGGCAGCAGACCTGTGGCCAAGTTCTTAGTGGCTGTGTCGTGTAG 1320
         |||||| |  ||||||||||||||||||||||||||||||| | |||| |||| |||
Db  1331 TGATGTGTGCTCAGGCAGCAGACCTGTGGCCAAGTTCTTAGTTGTGTCGTATGTCTCGTGGTAG 1390

Qy  1321 GACTGTAGAACTGTTAGAGGAAGAAACTGAACATTCCAGATGTGTGTAAATTGAATAAA 1380
         ||||||||        ||||||||||||||||||||||||  ||||| ||||  ||||
Db  1391 GACTGTAGA------AAGGGAACTGAACATTCCAGAGGCGTGTAGTGTAGTCAGTGAATCACGTAAA 1442

Qy  1381 GCTAGCCGTGATCCTTCAGCTGTTGCTGCATA 1411
         |||||     ||||| ||||||| | |||||
Db  1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1473
```

```
RESULT 14
ABZ42642
ID   ABZ42642 standard; DNA; 1679 BP.
XX
AC   ABZ42642;
XX
DT   04-MAR-2003  (first entry)
XX
DE   Human CXC chemokine receptor 4 nucleotide SEQ ID NO:75.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
KW   Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW   osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW   graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW   psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW   mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW   hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW   ulcer; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200261087-A2.
XX
PD   08-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US050107.
XX
PR   19-DEC-2000; 2000US-0257144P.
XX
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Burmer GC, Roush CL, Brown JP;
XX
DR   WPI; 2003-046718/04.
DR   P-PSDB; ABP81796.
XX
PT   New isolated antigenic peptides e.g., for G protein-coupled receptors
PT   (GPCR), useful for diagnosing and designing drugs for treating conditions
PT   in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT   autoimmune diseases.
XX
PS   Disclosure; Fig 1; 523pp; English.
XX
CC   The present invention describes antigenic peptides (I) comprising: (a)
CC   any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular G
CC   protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity or
CC   avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC   gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC   antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-related disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC   GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC   exemplification of the present invention
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
```

Query Match   55.1%; Score 933.4; DB 7; Length 1679;
Best Local Similarity 81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative   0; Mismatches 216; Indels 41; Gaps 5;

```
Qy         1 ATATACACTTCTGATAACTACACTCTGAGAAGTGGGTCTGGAGACTATGACTCCAACAAG 60
             ||||||||||||||||||||| ||||| ||||||||| ||| |||||||||| ||||||
Db       104 ATATACACTTCAGATAACTACACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163

Qy        61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC 120
             ||||||| ||||||| |||||||||||||| |||| ||| |||||||||| |||||||||
Db       164 GAACCCTGTTCCGTGAAGAAATGTCTAATTTCAATAAAATCTTCCTGCCCACCATCTAC 223

Qy       121 TTCATCATCTTCTTGACTGGCATGGCATGTCGGCAATGGATTGGTGATCCTCGTCATGGGTTAC 180
             |||||||||||  |||||||||||| ||||| ||||||| |||| |||||||||||||||||||
Db       224 TCCATCATCTTCTTAACTGGCATGTGGCCATGGATTGGTGATCCTTGGTCATGGGTTAC 283

Qy       181 CAGAAGAAGCTAAGGAGCAATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
             ||||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||
Db       284 CAGAAGAAACTGAGAAGCAATGACGGACAAGTACAGGCCGCACCTGTCAGTGGCCGACCTC 343

Qy       241 CTCTTTGTCATCACACTCCCTTCTCGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 300
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       344 CTCTTTGTCATCACGCTTCCCTTCTCGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG 403

Qy       301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
             |||| |||||||||| |||||||||| ||||||||||||||||||||||||||||||||
Db       404 AACTTCCTATGCAAGGCAGTCCATGTCATGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTC 463

Qy       361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCAACGGCCACCAACAGTCAA 420
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       464 ATCCTGGCCTTCATCAGTGCCTGGACCGGTACCTCGCCATCGTCCACGGCCACCAACAGTCAG 523

Qy       421 AGGCCAAGGAAACTGCTGCTCGGGCAGTAAAAGCAGTCTATGTGGGGTCTGGATCCAGCCCTC 480
             |||||||||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||
Db       524 AGGCCAAGGAAGCTGTTGGCTGAAAAGTGGTCTATGTTGGCGTCTGGATCCCTGCCCTC 583

Qy       481 CTCCTGACTATACCTGACTTCATCTTTGCGAGTCAGCGAGTCAGCAGGGGACATCAGTCAGGGG 540
             ||||||||||||||||||| |||||||||||||||||                   |||||||
Db       584 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG--------------------TCAGTGAGGCA 628

Qy       541 GATGACAGTGACATCTGTGACCGGCTTTACCCGACGATAGCCTGTGGATGGTGGTGTTCAA 600
             ||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||
Db       629 GATGACAGATATCTGTGACCGGTTCTACCCCAATGACTTGTGGTGTGTTGTTCCAG 688

Qy       601 TTCCAGCATATATAATGGTGGGTCCATCCTCGCCGGCATCGTCACTCCTCCGTTACTGC 660
             |||||| || ||| ||||||||||||||||||||||||||||||||||| ||||||||||
Db       689 TTTCAGCACATCATGGTGGCCTTATCCTGCTGGTATTGTCATCCGTCCTGCTCGTATTGC 748

Qy       661 ATCATCATCTAAGTGTCACACTCCAAGGGCCACCAGAGAGCGCAAGGCCCTCAAGACG 720
             ||||||||||| |||||||||||||||| ||||||||||||||||||||||||| ||||
Db       749 ATTATCATCTCCAAGTGTCACACTCCAAGGGGCCACCAGAGAGGGCAAGGCCCTCAAGACC 808

Qy       721 ACAGTCATCCTCATCCCAGCTTCTTGCCTGCTGGCTGCCATATGTGGGGGATCAGC 780
             |||||||||||||||| ||||||| ||||||| ||||||||||||||||||||||||||
Db       809 ACAGTCATCCTCATCCCGGCTTTCTTCTGCTGGCTGCCTTACTACCATTGGGATCAGC 868

Qy       781 ATCCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGAGATGTGACTTCGAGAGCATTGTG 840
             |||| ||||||||||||||||||| ||||||||||||||| |||||||||||| || |||||
Db       869 ATCCGCCTCCTCATCCTCCTGGAAATCATCAAGCAAGGTGTGAGTTTGAGAACACTGTG 928

Qy       841 CACAAGTGGATCCATCACAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCATC 900
             ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       929 CACAAGTGGATTTCATCACCGAGGGCCCCAGCTTTCTTCCACTGTTGCTGAACCCCATC 988

Qy       901 CTCTATGCCTTCCTGGGGCCAAGTTCAAAAGCTCTGCCCAGCGACTCAACTCCATG 960
             |||||||||||||| |||||||||| |||||||| ||||||||||||||||||||||
Db       989 CTCTATGCTTTCTTGAGCCAAATTAAAACCTCTCCCCAGCGACTCACCTCTGTG 1048

Qy       961 AGCAGAGGCTCCAGCTCAGATCCTTTCCAAAGGAAAGCGGGTGGACACTCTTCCGTC 1020
             ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db      1049 AGCAGAGGGTCCAGCCTCAGATCCTCCCAAAGGAAAGGAGGTGGACATTCATCTGTT 1108
```

ADC98645

ID   ADC98645 standard; cDNA; 1679 BP.
XX
AC   ADC98645;
XX
DT   01-JAN-2004   (first entry)
XX
DE   Human CXC chemokine receptor 4 (CXCR4)-encoding cDNA.
XX
KW   Human; CXC chemokine receptor 4; CXCR4; CXCR4 inhibitor;
KW   small cell lung cancer; SCLC; cellular adhesion; cellular proliferation;
KW   metastasis; motility; morphological change; drug screening; monitoring;
KW   cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   89..1147
FT   /tag= a
FT   /product= "Human CXCR4"
XX
PN   WO2003079020-A2.
XX
PD   25-SEP-2003.
XX
PF   20-MAR-2003; 2003WO-EP002916.
XX
PR   20-MAR-2002; 2002US-0366370P.
XX
PA   (DAND) DANA FARBER CANCER INST INC.
PA   (NOVS) NOVARTIS AG.
XX
PI   Salgia R;
XX
DR   WPI; 2003-757037/71.
DR   P-PSDB; ADC98646.
XX
PT   Use of a CXC chemokine receptor 4 (CXCR4) inhibitor, for modulating
PT   cellular adhesion or inhibiting proliferation, movement, or morphological
PT   change in a small cell lung cancer (SCLC) cell population, or for
PT   treating a subject having SCLC.
XX
PS   Disclosure; SEQ ID NO 1; 78pp; English.

XX
CC   The invention relates to methods and compositions for inhibiting small
CC   cell lung cancer (SCLC) proliferation and metastasis through modulation
CC   of the activity or expression of CXC chemokine receptor 4 (CXCR4), and
CC   optionally, that of the tyrosine kinase receptor c-Kit. Administration of
CC   a CXCR4 inhibitor modulates cellular adhesion and inhibits cellular
CC   proliferation, motility, and morphological changes in an SCLC cell
CC   population. The invention is based on the finding that CXCR4 is
CC   ubiquitously expressed, and c-Kit is variably expressed, in SCLC cells.
CC   The invention also discloses methods of screening for CXCR4 inhibitors;
CC   determining whether a sample of lung cancer cells expresses CXCR4 and is
CC   therefore amenable to CXCR4 inhibitor therapy; determining whether a
CC   sample of lung cancer cells coexpress CXCR4 and c-Kit is therefore
CC   amenable to treatment with a combination of a CXCR4 inhibitor and a c-Kit
CC   inhibitor; and a method of assessing whether CXCR4 inhibitor therapy or
CC   CXCR4 inhibitor/c-Kit inhibitor combination therapy should be continued,
CC   by determining CXCR4 (and optionally c-Kit) activity in two or more
CC   tumour samples taken from an SCLC patient during the course of treatment.
CC   The methods of the invention are useful for treating SCLC, for screening
CC   for agents for the treatment of SCLC, for determining whether a SCLC from
CC   individual patients is suitable for treatment with agents of the
CC   invention, and for determining whether such treatment is effective and
CC   should be continued. The present sequence represents cDNA encoding human
CC   CXCR4, which can be used in methods of the invention.
XX
SQ   Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
XX

Query Match   55.1%; Score 933.4; DB 9; Length 1679;
Best Local Similarity   81.8%;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

```
Qy    1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACATGACTCCAACAAG 60
         |||||||||| ||||||||||    ||||| |   || ||| ||| ||||||||||  ||
Db  104 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
         |||||||| ||||| ||||| ||| | ||||||||| |||||||||||||||||||||||
Db  164 GAACCCTGTTCCGTGAAGAAAATGCTAATTCAATAAAAATCTTCCTGCCACCATCTAC 223

Qy  121 TTCATCATCTTCTGACTGCCATAGTCGGCAATGATTGGTGATCCTGGTCATGGTTAC 180
         | ||||||||||| |||| ||||||| | |||||||||||| ||| ||||||||||||
Db  224 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGATTGGTCATCCTGGTCATGGTTAC 283

Qy  181 CAGAGAAGCTAAGGAGCAATGACCGGCAAGTACCGGCTCAGTGGCTCAGTGCGCCACCTC 240
         |||||||| ||| || ||||||| | |||||| || ||||||||||||||||||||||||
Db  284 CAGAGAAACTGAGAAGCAATGACGGACAAGCAAGTACGGCTCAGTGGCTCAGTGCGCCACCTC 343

Qy  241 CTCTTTGTCATCACACTCCCCTTCGGGCCAGTTGATGCCATGGCGACTGGTACTTTGGG 300
         ||||| ||||||||| |||||||| || |||| |||| |||| || |||||||||||||
Db  344 CTCTTTGTCATCACGCTTCCCTTCTGGCCAGTTGATGCCAGTGCAAACTGGTACTTTGGG 403

Qy  301 AAATTTTGTGTAAGGCTGTCCATATCATCTCACACTCTACACCTCTACAGCAGCGTTCTC 360
         || || |||| |||||||||||| | ||| ||||| |||||| ||||||||||||||||
Db  404 AACTTCCTATGCAAGGCAGTCCATGTCATCTCACACTCTACACCTCTACAGCAGTGTCCTC 463

Qy  361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTGCCATTGTCACGCGCCACCAACAGTCAA 420
         |||||||||||||||||||||| | ||||||| |||| |  ||||||||||||| |||||
Db  464 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCCATCGTCCACGCGCCACCAACAGTCAG 523

Qy  421 AGGCCAAGGAGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGATCCCAGCCCTC 480
         |||||||||| |||||||||||||||||||| ||| |||| ||| |||||||||||||| |
Db  524 AGGCCAAGGAGCTGCTGGCTGAAAAGGTGTTCATGTTGGCGTCTGATGTTGGGCTCATGGG 583

Qy  481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
         |||||||||| ||||||||||||||||| || |||       ||||||||||||||||
Db  584 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG----------TCAGTGAGGCA 628

Qy  541 GATGACAGGTACATCTGTGACCGCCTTTACCCGATAGCCTGTGGATGGTGGTGTTTCAA 600
         |||||| ||| | || |||| |||||| |||||||||| | |||||| ||||| ||||
Db  629 GATGACAGATATCTGTGACCGCCTTCACCCCCAATGACTTGTGGGCGTGTGTTCCAG 688

Qy  601 TTCCAGCATATAATGGGTGGTCTCATCCTGCCCGGCATGCTCATCCTCTCGTGTTACTGC 660
         ||||| |||||||||||||||| |||||||||||||||| |||||||||||||||||||
Qy 1021 TCCACGGAGTCAGAATCCTCCAGCTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
         |||| | |||| |  ||| ||||| ||| |||||||||||||       || ||||| ||
Db 1109 TCCACTGAGTCGAGTCTTCAAGTTTTCACTCCAGCTAA-------CACAGATGTAAAA 1160

Qy 1081 AATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAGAG 1140
         | | | || || || | |||| |||||||||||||||||||||| |||||||||||
Db 1161 GACTTTTTTTTATACGATAAATAACTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1220

Qy 1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
         ||||||||  | ||||| ||||||||          |||||| ||||  |||| ||||| |
Db 1221 ACTGACCAATATTGTACAGTTTTTTATTGCTTGTGTTGGATTTTGT------CTTGTGTTTC 1274

Qy 1201 TCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAAATATTGTTTTTGTTGTTCATGT 1260
         | ||||||||||||| |||||  |||||| ||| ||| ||||    ||||||||||| |
Db 1275 TTTAGTTTTTGTGAAGTTTGAAGTTTAATTGACTTATTTATA----TAAATTTTTTTGTTTTCATAT 1330

Qy 1261 GAATGAGCCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATTCTGTGTAG 1320
         | ||| | |||||||||||||||||||||||||||||| || |||| | || ||| |||||
Db 1331 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTCAGTTGCTCGTATGTCTCGTGGTAG 1390

Qy 1321 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAAATGTGGTAAATTGAATAAA 1380
         |||||||||||        |||| |||||||||||||||   |||| |||||||||| |||
Db 1391 GACTGTAGAA------AAGGGAACTGAACATTCCAGAGCCGTGTAGTGAATCACGTAAA 1442

Qy 1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
         ||||| | |||||| ||||||||| ||||||
Db 1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1473
```

```
Db   689  TTTCAGCACATCATGGTTGGCCTTATCCTGCCTGGTATTGTCATCCTGTCCTGCTATTGC  748

Qy   661  ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGGCCAAGGCCCTCAAGACG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   749  ATTATCATCTCCAAGCTGTCACACTCCAAGGGGCCACCAGAAGGCCAAGGCCCTCAAGACC  808

Qy   721  ACAGTCATCCTCATCCTAGCTTCTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGC  780
          ||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   809  ACAGTCATCCTCATCCTCGGCTTTCTTCGCCTGTTGGCTGCCTTACTACATTGGGATCAGC  868

Qy   781  ATCGACTCCCTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAAGAGCATTGTG  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   869  ATCGACTCCCTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG  928

Qy   841  CACAAGTGGATCTCCATCACAGAGCCCTCGCCTTCTTCCACTGTGCCTGAACCCCATC  900
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   929  CACAAGTGGATTTCCATCACCGAGGGCCCTAGCTTCTTCTTCCACTGTTGCTGAACCCCATC  988

Qy   901  CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG  960
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   989  CTCTATGCCTTCCTTGGAGCCAAATTTAAAACCTCTGCCCAGCACGCACTCACCTCTGTG  1048

Qy   961  AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC  1020
          ||||||||||||||||||||||||||||||||||||||||||||    ||    ||||
Db   1049 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTT  1108

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT  1080
          |||||||||||||||||||||||||||||||||||||||||    ||  ||| |
Db   1109 TCCACTGAGTCTGAGTCTTCAAGTTTTCACTCCAGCTAA-------CACAGATGTAAAA  1160

Qy   1081 AATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG  1140
          |||||||||||||||| |||||||||||||||||||||||
Db   1161 GACTTTTTTTTTATACGATAAATAACTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG  1220

Qy   1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTANGTTCC  1200
          |||||||||||||||||||||||||| | ||| |||| |||| |||
Db   1221 ACTGACCAATATTGTACAGTTTTTATTGCTTGGATTTTTGT------CTTGTGTTTC  1274

Qy   1201 TCTAGTTTTTGTGAGGTTTGACTTAATTATATAAAATATTGTTTTGTTTTGTTTCATGT  1260
          |||||||||||||| |||||| ||| ||||| |||||||  ||||||||||
Db   1275 TTTAGTTTTTGTGAAGTTTAATTGACTTATTTATA---TAAATTTTTTTTGTTTCATAT  1330

Qy   1261 GAATGAGCCTCTAGGCAGCAGACCCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTAG  1320
          |||||||||||||||||||||||||||||||||||| ||| ||||||||||
Db   1331 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTGTATGTCTCGTGGTAG  1390

Qy   1321 GACTTGTAGAACTGTAGAGGAAGAAGAAACTGAACATTCCAGAATGTGTGGCTAAAATTGAATAAA  1380
          ||||||||||   |||    |  |||||||||||||||||||||  |||||||||||
Db   1391 GACTTGTAGAA-------AAGGGAACTGAACATTCCAGAGCGTGTGAGTGAATCACGTAAA  1442

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA  1411
          |||||  |||||||| ||||||
Db   1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA  1473
```

Search completed: May 17, 2004, 06:37:17
Job time : 736.44 secs

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      May 17, 2004, 05:52:09 ; Search time 135.63 Seconds
                                       (without alignments)
                                       6931.281 Million cell updates/sec

Title:       US-09-367-052-7
Perfect score: 1694
Sequence:    1 atatacacttcgtaacta........taaaagtcaatgataaactt 1694

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched:    682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:      1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents NA:*
             1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
             2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
             3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
             4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
             5: /cgn2_6/ptodata/2/ina/pcTUS_COMB.seq:*
             6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

             Pred. No. is the number of results predicted by chance to have a
             score greater than or equal to the score of the result being printed,
             and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	933.4	55.1	1664	4	US-09-582-224A-5	Sequence 5, Appli
2	933.4	55.1	1664	4	US-09-023-655-1213	Sequence 1213, Ap
3	933.4	55.1	1679	4	US-09-517-605-14	Sequence 14, Appl
4	933.4	55.1	1737	1	US-08-202-056-4	Sequence 4, Appli
5	933.4	55.1	1737	1	US-08-076-093A-3	Sequence 3, Appli
6	933.4	55.1	1737	2	US-08-701-265-3	Sequence 3, Appli
7	933.4	55.1	1737	2	US-08-284-586-3	Sequence 3, Appli
8	933.4	55.1	1737	2	US-08-805-478-3	Sequence 3, Appli
9	933.4	55.1	1737	2	US-08-802-627A-3	Sequence 3, Appli
10	933.4	55.1	1737	2	US-08-801-238-3	Sequence 3, Appli
11	933.4	55.1	1737	2	US-08-801-228-3	Sequence 3, Appli
12	933.4	55.1	1737	2	US-09-104-296-3	Sequence 2, Appli
13	933.4	55.1	1737	5	PCT-US94-06380-2	Sequence 2, Appli
14	855.6	50.5	1225	4	US-09-016-434-1235	Sequence 1235, Ap
15	834	49.2	1317	3	US-08-153-848-45	Sequence 45, Appl
16	834	49.2	1317	4	US-09-299-843A-45	Sequence 45, Appl
17	834	49.2	1317	4	US-09-088-337B-45	Sequence 45, Appl
18	834	49.2	1317	5	PCT-US93-11153-45	Sequence 45, Appl
19	834	11.7	1119	1	US-09-170-496D-65	Sequence 65, Appl
20	198.4	11.7	1679	1	US-08-202-056-6	Sequence 6, Appli
21	198.4	11.7	1679	1	US-08-076-093A-5	Sequence 5, Appli
22	198.4	11.7	1679	1	US-08-701-265-5	Sequence 5, Appli
23	198.4	11.7	1679	2	US-08-284-586-5	Sequence 5, Appli
24	198.4	11.7	1679	2	US-08-805-478-5	Sequence 5, Appli
25	198.4	11.7	1679	2	US-08-802-627A-5	Sequence 5, Appli
26	198.4	11.7	1679	2	US-08-801-238-5	Sequence 5, Appli
27	198.4	11.7	1679	2	US-08-801-228-5	Sequence 5, Appli
28	198.4	11.7	1679	3	US-09-104-296-5	Sequence 5, Appli
29	198.4	11.7	1679	5	PCT-US94-06380-3	Sequence 3, Appli
30	198.4	11.7	2818	3	US-08-982-493-7	Sequence 7, Appli
31	198.4	11.7	2818	3	US-08-628-655-1	Sequence 1, Appli
32	196.8	11.6	1119	4	US-09-170-496D-199	Sequence 199, App
33	195.8	11.6	2085	4	US-09-299-843A-65	Sequence 65, Appl
34	195.8	11.6	2085	4	US-09-088-337B-65	Sequence 65, Appl
35	193.8	11.4	1107	4	US-09-170-496D-19	Sequence 19, Appl
36	193.8	11.4	1293	4	US-09-016-434-1052	Sequence 1052, Ap
37	193.8	11.4	1670	3	US-08-709-838-1	Sequence 1, Appli
38	193.8	11.4	1670	3	US-08-829-839-1	Sequence 1, Appli
39	193.8	11.4	1670	4	US-09-023-655-980	Sequence 980, App
40	193.8	11.4	1670	4	US-09-624-594-1	Sequence 1, Appli
41	192.2	11.3	1107	1	US-09-170-496D-173	Sequence 173, App
42	186.8	11.0	2751	1	US-08-153-848-23	Sequence 23, Appl
43	186.8	11.0	2751	3	US-09-299-843A-23	Sequence 23, Appl
44	186.8	11.0	2751	4	US-09-088-337B-23	Sequence 23, Appl
45	186.8	11.0	2751	5	PCT-US93-11153-23	Sequence 23, Appl

ALIGNMENTS

```
RESULT 1
US-09-582-224A-5
; Sequence 5, Application US/09582224A
; Patent No. 6429908
; GENERAL INFORMATION:
;  APPLICANT: IIJIMA, Osamu
;  APPLICANT: GOTO, Takeshi
;  APPLICANT: SHIMADA,Takashi
;  TITLE OF INVENTION: HIV Infection Inhibitors
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/09/582,224A
;  CURRENT FILING DATE: 2000-07-21
;  PRIOR APPLICATION NUMBER: PCT/JP99/06534
;  PRIOR FILING DATE: 1999/11/24
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: Microsoft Word
; SEQ ID NO 5
;  LENGTH: 1664
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: cDNA of CXCR4
US-09-582-224A-5

Query Match          55.1%;   Score 933.4;  DB 4;  Length 1664;
Best Local Similarity 81.8%;  Pred. No. 1e-210;
Matches 1154;  Mismatches 216;  Conservative 0;  Indels 41;  Gaps 5;

Qy    1 ATATACACTTCGTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
          ||||||||||||| ||||||||| || |||||||||||||| ||||||||||||||||||
Db   92 ATATACACTTCGATAACTAACTACCGAGGAAATGGGCTCAGGGACTATGACTCCATGAAG 151

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
         ||||||| |||||  || || |||  ||  ||||||||  |||||||| ||||||||||
Db  152 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC 211

Qy  121 TTCATCATCTTCTGACTGGCATAGTCGGCAATGGCTAGTCCGGTCATGGGTAC 180
         | ||||||||||| || |||||| | ||| ||||| | |||||||||||||||
Db  212 TCCATCATCTTCTTAACTGGCATGGCATGGCAATGGATTGGTCATCCTGGTCATGGGTAC 271

Qy  181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTCAGTGGCGTGACCCTC 240
         ||||||||  | || |||||||||||||||||||||||| || ||||| |||||
Db  272 CAGAAGAAACTGAGAAGCATGACGGACAAGTACCGGCTCACCTGTCAGTGGCCGACCCTC 331

Qy  241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGTGACTGGTACTTTGGG 300
         ||||||||||||| || || |||||||||||||||||| ||||   |  |||||||||||
Db  332 CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG 391

Qy  301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
         |||||||||||  || ||||||||||||||| |||||||||||||||||||||| ||||
Db    ...
```

RESULT 2
US-09-023-655-1213

; Sequence 1213, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/023,655
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1213:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1664 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g219868
;US-09-023-655-1213

Query Match   55.1%;   Score 933.4;   DB 4;   Length 1664;
Best Local Similarity 81.8%; Pred. No. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
            |||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db   92 ATATACACTTCAGATAACTACTACACCGAGAAATGGGCTCAGGGGACTATGACTCCATGAAG 151

Qy   61 GAACCCTGCTTCCGGGATGAAAAACGTCATTTCAATTGAATAGGATCTTCCTGCCACCATCTAC 120
            |||||||||||||||||||||||| || ||||||| ||||||| |||||||||||||||||||||||
Db   152 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCCACCATCTAC 211

Qy   121 TTCATCATCTTCTTGACTCGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
            |||||||||||| ||||||||||||| ||||||| |||||||| ||||||||||||||||||
Db   212 TCCATCATCTTCTTAACTCGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 271

Qy   181 CAGAAGAAGCTAAGGACGTCAAGCAACAGCTACCGGCTGCACCTGTCAGTGGGCTGACCTC 240
            |||||||||||||||||| ||||||||||||||| |||||||||||||||||||||||||||||
Db   272 CAGAAGAAACTGAGAAGCATGACGGACACAGTAGACGTGGCCACCTC 331

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGCCGTTCGAGTGCATGCATGCCATGGTACTTTGGG 300
            ||| |||||||||| |||||||||||||| |||| |||||||||||||||||||||

```
Qy   1381 GCTAGCCGTGATCCTCAGCTGTGCTGCATA 1411
          ||||  ||||||||||||||||| ||||||
Db   1431 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1461

RESULT 3
US-09-517-605-14
; Sequence 14, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
;   APPLICANT: Littman, Dan R.
;   APPLICANT: Kwon, Douglas S.
;   APPLICANT: van Kooyk, Yvette
;   APPLICANT: Geijtenbeck, Tneo
;   TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
;   TITLE OF INVENTION: CELLS
;   FILE REFERENCE: 1049-1-017
;   CURRENT APPLICATION NUMBER: US/09/517,605
;   CURRENT FILING DATE: 2000-03-02
;   NUMBER OF SEQ ID NOS: 17
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-517-605-14

  Query Match       55.1%;  Score 933.4;  DB 4;  Length 1679;
  Best Local Similarity   81.8%;  Pred. No. 1e-210;
  Matches 1154;  Conservative    0;  Mismatches 216;  Indels   41;  Gaps    5;
```

```
          ZIP:  94080
COMPUTER READABLE FORM:
   MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
   COMPUTER:  IBM PC compatible
   OPERATING SYSTEM:  PC-DOS/MS-DOS
   SOFTWARE:  patin (Genentech)
CURRENT APPLICATION DATA:
   APPLICATION NUMBER:  US/08/202,056
   FILING DATE:  25-FEB-1994
   CLASSIFICATION:  436
PRIOR APPLICATION DATA:
   APPLICATION NUMBER:  07/677211
   FILING DATE:  29-MAR-1991
ATTORNEY/AGENT INFORMATION:
   NAME:  Love, Richard B.
   REGISTRATION NUMBER:  34,659
   REFERENCE/DOCKET NUMBER:  706P3
TELECOMMUNICATION INFORMATION:
   TELEPHONE:  415/225-5530
   TELEFAX:  415/952-9881
   TELEX:  910/371-7168
INFORMATION FOR SEQ ID NO:  4:
SEQUENCE CHARACTERISTICS:
   LENGTH:  1737 bases
   TYPE:  nucleic acid
   STRANDEDNESS:  single
   TOPOLOGY:  linear
US-08-202-056-4
```

```
Query Match      55.1%;    Score 933.4;  DB 1;   Length 1737;
Best Local Similarity  81.8%;  Pred. No. 1e-210;
Matches 1154;  Conservative  0;  Mismatches 216;  Indels 41;  Gaps  5;

Qy     1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAG    60
          ||||||||||||| ||||| ||||||| |||||||||  ||||||||||||||||||||||
Db   106 ATATACACTTCAGATAACTACTCACCGAGGAAATGGGCTCAGGGACCATGACTCCATGAAG   165

Qy    61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC   120
          |||||||||||||||||||  ||||||||||||||||  ||||| |||||||||||||||
Db   166 GAACCCCTGTTTCCGTGAAGAAATGCTAATTCAATAAAATCTTCCTGCCCACCATCTAC   225

Qy   121 TTTCATCTTCTTGACTGGCATAGTCGGCAATGATGGTTGTCATCCTGGTCATGGGTTAC   180
          ||||||||||||||||||  ||||| ||||  ||||||||||||||||||| |||||||||
Db   226 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGATTGGTCATCCTGGTCATGGGTTAC   285

Qy   181 CAGAAGAAGCTAGGAGCCATGACGACGACAAGTACCGGCTCACCTGTGACCTC   240
          |||||||| |||||||||||||||||  |||||| |||||||||||||||||||
Db   286 CAGAAGAAACTGAGAAGCATGACGACGACAAGTACAAGGCTCACCTGTGACCTC   345

Qy   241 CTCTTGTCATCACACTCCCCTTCTGGCAGTTGATGCAGCCATGGTGATGCCGGTGCAAACTGGTACTTTGGG   300
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 CTCTTTGTCATCACGCTTCCCTTCTGGCAGTTGATGCCGGTGCAAACTGGTACTTTGGG   405

Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATCTCACACTGTCAACCTCACACTCACAGTGTCCTC   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 AACTTCCTATGCAAGGACGCAGTCCATGTCATCTCACACTGTCAACAACAGTCAG   465

Qy   361 ATCCTGGCCTTCATCAGCTGGACCGGTACCTGCCATTGTCCAACAGTCAA   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466 ATCCTGGCCTTCATCAGTCTCGGACCGTACCTGGCCATCGTCCAGCCACCAACAGTCAG   525

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGGCGTCTGGATCCCAGCCCTC   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 AGCCAAGGAAGCGAGCTGTTGGCTGAAAAGCGTGGTCTATGTGGGTCTGAAATCCCTGCCCTC   585

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTGCAGCCAGGGGGACATCAGCAGTCAGGGG   540
          ||||||||||||||||||||||||||||||||||||| ||||||||
Db   586 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG---------TCAGTGAGGCA   630

Qy   541 GATGACAGGTACATCTGTGACCGCCGTTTACCCGATGGTCGTGTTCAA   600
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   631 GATGACAGATATATCTGTGACCGGTTCACCCCAATGACTTGTGGGGTGTGTTCCAG   690
```

```
Db   689 TTTCAGCACACATCATGGTTGGCCTTATCCTGCCTGGTATTGTCATCCTGTCCTGCTATTGC   748
          |||||||||||
Qy   661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTTCAAGACG   720
          ||||||||||||
Db   749 ATTATACATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAAGCCGCAAGGCCCTTCAAGACC   808
          ||||||||||||
Qy   721 ACAGTCATCCTCCAGCTTCTTGCCTGCTGGCTGGTGCCATAATTATGTGGGGATCAGC   780
          ||||||||||||
Db   809 ACAGTCATCCTCCATCCTGGCTTTCTTGCGCCTGCTGGCTGGTGCCATAATTATGTGGGGATCAGC   868
          ||||||||||||
Qy   781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAGGATGTGACTTCGAGAGCATTGTG   840
          ||||||||||||
Db   869 ATCGACTCCTTCATCCTTTGGGAAATCATCAAGCAGGGGTGTGAGTTTGAGAACACTGTG   928
          ||||||||||||
Qy   841 CACAAGTGGATCTCCATCACAGAGAGCGCCCTCCCCTTCTTCCACTGTGCCTGAACCCCATC   900
          ||||||||||||
Db   929 CACAAGTGGATTTCCATCACCGAGCCCTCAGCTTTCTTCCACTGTGTCTGAACCCCATC   988
          ||||||||||||
Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG   960
          ||||||||||||
Db   989 CTCTATGCCTTTCCTGGAGCCAAATTAAAACCTCTGCCCAGCACGCACTCACCTCTGTG   1048
          ||||||||||||
Qy   961 AGCAGAGGCTCCAGCTCAAGATCCTTCCAAAGGAAAGGCGGGTGGCACACTCTTCCGTC   1020
          ||||||||||||
Db  1049 AGCAGAGGGTCAGCCTCAAGATCCTCTCCAAAGGAAAGGCGAGGTGGCATTCATCTGTT   1108
          ||||||||||||
Qy  1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTGTTCACTCCAGTGTAACCCTTATGCAAAGACTTATAT   1080
          ||||||||||||
Db  1109 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAA-------CACAGATGTAAAA   1160
          ||||||||||||
Qy  1081 AATATATATATATATATGATAAGAACTTTTTATGTTACACATTTTCAGATATAAAAG   1140
          ||||||||||||
Db  1161 GACTTTTTTTTTTATACGATAAATAACTTTTTTTAAGTTACACATTTTCAGATATAAAAG   1220
          ||||||||||||
Qy  1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC   1200
          ||||||||||||
Db  1221 ACTGACCAATATTGTACAGTTTTTTATTGCTTGGTTGT------CTTGTGTTTC   1274
          ||||||||||||
Qy  1201 TCTAGTTTTTTGTGAGGTTTGACTTTAATTATATAAAATATTGTTTTTTGTTTGTTTCATGT   1260
          ||||||||||||
Db  1275 TTTAGTTTTTGTGAAGTTTTGAAGTTTAATTGACTTATTTATA----TAAATTTTTTTGTTTCATAT   1330
          ||||||||||||
Qy  1261 GAATGAGCGTCTAGGCAGCAGAACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTAG   1320
          ||||||||||||
Db  1331 TGAATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGGTAG   1390
          ||||||||||||
Qy  1321 GACTGTAGAACTGTGAGAGAAGAAGAAACTGAACATTCCAGAAATGTGTGGTAAATTGAATAAA   1380
          ||||||||||||
Db  1391 GACTGTAGAA-------AAGGGAACTGAACATTCCAGAGCGGTGTAGTGAATCACGTAAA   1442
          ||||||||||||
Qy  1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA   1411
          ||||||||||||
Db  1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA   1473
```

```
RESULT 4
US-08-202-056-4
; Sequence 4, Application US/08202056
; Patent No. 5440021
GENERAL INFORMATION:
   APPLICANT:  Chuntharapai, Anan
   APPLICANT:  Hebert, Caroline
   APPLICANT:  Kim, Kyung Jin
   APPLICANT:  Lee, James
   TITLE OF INVENTION:  Antibodies to Human IL-8 Type B Receptor
   NUMBER OF SEQUENCES:  8
   CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Genentech, Inc.
      STREET:  460 Point San Bruno Blvd
      CITY:  South San Francisco
      STATE:  California
      COUNTRY:  USA
```

```
          COUNTRY: USA
          ZIP: 94080
     COMPUTER READABLE FORM:
          MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: WinPatin (Genentech)
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/08/076,093A
          FILING DATE: 11-Jun-1993
          CLASSIFICATION: 530
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 07/810782
          FILING DATE: 19-DEC-1991
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 07/677211
          FILING DATE: 29-MAR-1991
     ATTORNEY/AGENT INFORMATION:
          NAME: Love, Richard B
          REGISTRATION NUMBER: 34,659
          REFERENCE/DOCKET NUMBER: 706P2
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: 415/225-5530
          TELEFAX: 415/952-9881
          TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 1737 nucleotides
          TYPE: Nucleic Acid
          STRANDEDNESS: Single
          TOPOLOGY: Linear
US-08-076-093A-3

Query Match        55.1%;  Score 933.4;  DB 1;  Length 1737;
Best Local Similarity 81.8%;  Pred. No. 1e-210;
Matches 1154;  Conservative    0;  Mismatches 216;  Indels  41;  Gaps  5;

Qy    1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
         |||||||| ||||||||| || |||||||||| |||||  |||  |||||||||||||||
Db  106 ATATACACTTCAGATAACTACACACCGAGAAAATGGGCTCAGGGGACTCATGACTCCATGAAG 165

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATTGGCCAATGAATGGATTGGTCATGGTTAC 120
         |||||||| ||||| |||||||| ||  |||||||| |||| |  ||   |||| ||| ||  ||
Db  166 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCCACCATCTAC 225

Qy  121 TTCATCATCTTCTGACTGGCATAGTCGGCAATGAATGGATTGGTGATCGTCGTCATGGGTTAC 180
         | ||||||||||| |||||| || ||| |||||||||||||||| |||  || ||||| ||||
Db  226 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTGTCATCCTGGTCATGGGTTAC 285

Qy  181 CAGAGAGCTAAGGAGCATAGGAGGACGGACGATGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
         ||||||  | ||||  || || |||  || ||| ||| ||||||  || |||||||||||||||||  |||||
Db  286 CAGAGAAACTGAGAAACTGAGACAGATGACGAGCAAGTACCAGTGCACCTGTCAGTGGCTGGCCGACCTC 345

Qy  241 CTCTTTGTCATCACACTCCCCTTCTCGGCAGTTGATGCCATGGTCGTGACTGCTGACTTTGG 300
         |||||||||||||  | ||||| || ||||||||||||  ||| |  |   || ||||| ||
Db  346 CTCTTTGTCATCAGCGCTTCCCTTCCCTTCGGCAGTTGATGCGTGGCAAACTGGTACTTTGG 405

Qy  301 AAATTTTGTGTAAGGCTGTCCATATCATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
         || ||| || |||||| |||| ||| |||||||| ||||||||||||||||||||| | |||
Db  406 AACTTCCATGCAAGGCAGTCATGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTC 465

Qy  361 ATCCTGGCCTTCATCAGCCTGGACCCGGTACCTCGCCATTGTCCAGCGCCACCAACAGTCAA 420
         |||||||||||||||||| ||||||||  | | |||| ||||| | |||| |||||||||| |
Db  466 ATCCTGGCCTTCATCAGTCTGGACCCGCCATCGCCACCTGCCATCGTCACGCCACCAACAGTCAG 525

Qy  421 AGGCCAAGGAAAACTGCTGGCTGGAAAAGGCAGTCTATGTGGGGCTCTGGATCCCAGCCCTC 480
         ||||||||||| ||| |||||||| |||| |||| ||||| ||| ||||||||||| |||||
Db  526 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTTGGGGTCTGGATCCCGCCCTC 585

Qy  481 CTCCTGACTATACCTGACTTCATCTTTGCCGACATACTGCTGGGG 540
         || |||||||| ||||||||||||||||||| ||          |
Db  586 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG--------TCAGTGAGGCA 630

Qy  601 TTCCAGCATAATAATGGTGGTCTCATCTGCCCGGCATCGTCATCCTCCTCGTTACTGC 660
         || |||||| |  | ||| ||  | ||||| | ||  |||||||| ||| ||| ||||
Db  691 TTTCAGCACATCAGATGGTTGGCCTTATCCTGCCTGGTATTGTCATCCTGCCGCTATTGC 750

Qy  661 ATCATCATCTCTAAGCTGTCACACTCCAAGTGGTCGCCACCCAAGAGGGCCACCAGAAGGCCCTCAAGACG 720
         || ||||||||  |||||||||||| | |         || |||||||||| ||  ||||| |||||||| |
Db  751 ATTATCATCTCCAGCTGTCACACTGTCACACTCCAAGAGGGCCAAGGACCCTCAAGACC 810

Qy  721 ACAGTCATCCTCCATCCTAGCTTCTTCTTGCTGCTGCCATATTGTGGGGATCAGC 780
         ||||||||||||||||||| ||||||| |||| |||  ||| ||| ||| ||||
Db  811 ACAGTCATCCTCCATCCTCAGCTTCTTCGCCCGTGTGGCTGCCTTACTACATTGGGATCAGC 870

Qy  781 ATGGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
         |||||||||||||||| |||| |||||||||||||||  |||||| |||||| |||||
Db  871 ATGGACTCCTTCATCCTCTGGAAATCATCAAGCAAGGAGGGTGTGAGTTTGAGAACACTGTG 930

Qy  841 CACAAGTGGATCTCCATCACAGAGAGGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATC 900
         ||||||||| | |||||||| ||| |||||| |  ||||||| ||| ||||||||||||||
Db  931 CACAAGTGATTTCCATCACCGAGGGCCCTAGCTTCTTCCTGCTGTTCTGAACCCCATC 990

Qy  901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCGCCCAGCATGCACTCAACTCCATG 960
         |||||||| |||||| || |||||  |||||| |||| |||||| |||||| |||  | ||||
Db  991 CTCTATGCTTTCCTTGGAGCCCAAATTTAAAAACCTCTGCCCAGCCAGCACGCACCTCTGTG 1050

Qy  961 AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAGGCGGGGGGTGGACACTCTTCCGTC 1020
         ||||||| ||||||||||||||||| |||||||||||| || ||  |||||| | ||| || 
Db 1051 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAGCGAGGTGGACATTCATCTGTT 1110

Qy 1021 TCCACGGGAGTCAGAATCCTCCCAGTTTCACTCCAGCCTAACCCCTTATGCAAAGACTTATAT 1080
         |||||| || ||    | ||   ||||||||||||||  |         | |  |||||||| 
Db 1111 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAA------CACAGATGTAAA 1162

Qy 1081 AATATATATATATATATGATAAGAACTTTTTTTATGTTACACATTTCCAGATATAAGAG 1140
             | ||    |||||| ||||| |||||||||   |||| ||||||||| ||||||| ||
Db 1163 GACTTTTTTTTTATACGATAAATAACTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1222

Qy 1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
         |||||||| | |||||||||||||| |||| ||  | |||||||||||    ||| |||
Db 1223 ACTGACCAATATTGTACAGTTTTTTTATTGCTTGTTGGGATTTTTGT------CTTGTGTTTC 1276

Qy 1201 TCTAGTTTTTGTGAGGTTTGACTTAATTTATAATATAAATATGTTTTTGTTTGTTTCATGT 1260
         | |||||||||||| ||| | |||| ||||| ||    ||||  ||||||||||||||| |
Db 1277 TTTAGTTTTTGTGAAGTTTAATTGACTTATTATA----TAAATTTTTTTTTTGTTTCATAT 1332

Qy 1261 GAATGAGCGTCTAGGCAGCAGACCTGTGGCCAAGTCTTAGTAGCTGTTTATCGTGTGTAG 1320
         ||| | | ||||||||| ||||||||||||||||| ||||| |||||| | | | ||| ||||
Db 1333 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTGTATGTCTCGTGGTAG 1392

Qy 1321 GACTGTAGAACTGTAGAGGAAGAAGAAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
         |||||||||         ||  |||  || ||  |      | ||  | ||| ||| || |||
Db 1393 GACTGTAGAA-------AGGGGACTGAGAA-AGGGGACTGAGAATCACGTAAA 1444

Qy 1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
         ||| |   ||||||  ||||||| | |||||
Db 1445 GCTGAAAATGATCCCCAGCTGTTTATGCATA 1475
```

RESULT 5
US-08-076-093A-3
; Sequence 3, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:
;   APPLICANT: Chuntharapai, Anan
;   APPLICANT: Lee, James
;   APPLICANT: Hebert, Caroline
;   APPLICANT: Jin Kim, K.
;   TITLE OF INVENTION: Antibodies to Human PF4A Receptors
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California

```
        ADDRESSEE:  Genentech, Inc.
        STREET:   460 Point San Bruno Blvd
        CITY:   South San Francisco
        STATE:   California
        COUNTRY: USA
        ZIP:  94080
  COMPUTER READABLE FORM:
        MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
        COMPUTER:  IBM PC compatible
        OPERATING SYSTEM:   PC-DOS/MS-DOS
        SOFTWARE:  WinPatin (Genentech)
  CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/701,265
        FILING DATE:  22-AUG-1996
        CLASSIFICATION:  424
  PRIOR APPLICATION DATA:
        APPLICATION NUMBER:  08/076093
        FILING DATE:   11-Jun-1993
        APPLICATION NUMBER:  07/810782
        FILING DATE:  19-DEC-1991
  PRIOR APPLICATION DATA:
        APPLICATION NUMBER:  07/677211
        FILING DATE:  29-MAR-1991
  ATTORNEY/AGENT INFORMATION:
        NAME:  Love, Richard B
        REGISTRATION NUMBER:  34,659
        REFERENCE/DOCKET NUMBER:  706P2
  TELECOMMUNICATION INFORMATION:
        TELEPHONE:  415/225-5530
        TELEFAX:  415/952-9881
        TELEX:  910/371-7168
  INFORMATION FOR SEQ ID NO:   3:
    SEQUENCE CHARACTERISTICS:
        LENGTH:  1737 nucleotides
        TYPE:  Nucleic Acid
        STRANDEDNESS:  Single
        TOPOLOGY:  Linear
US-08-701-265-3
```

```
Query Match                 55.1%;     Score 933.4;  DB 1;   Length 1737;
Best Local Similarity     81.8%;     Pred. No. 1e-210;
Matches 1154;   Conservative    0;  Mismatches 216;   Indels  41;   Gaps   5;

Qy      1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAG 60
            ||||||||||||||||||| ||||||||||||||||||||||| |||||||||||||||
Db    106 ATATACACTTCAGATAACTACTACACCGGAGGAAATGGGGTCTCAGGGGACTATGACTCCCATGAAG 165

Qy     61 GAACCCTGCTTCCTGGGATGAAAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 120
            ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    166 GAACCCTGTTTCCTGAAGAAAATGTAATTTCAATAAAATCTTCCTGCCACCATCTAC 225

Qy    121 TTCATCATCTTCTTGACTGGCATAGTGGCAAGTGGTGATCCTGGTCATGGGTTAC 180
            ||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    226 TCCATCATCTTCTTAACTGGCCAATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 285

Qy    181 CAGAAGAAGCTAAGGAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCGACCTC 240
            ||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    286 CAGAAGAAACTGAGAAGCATGACGGACAAGTACCAGGCTACAGGCTGCACACCTGTCCGACCTC 345

Db    346 CTCTTTGTCATCACACCCGTTCCCTTCTGGGCAGTTGATGCAGTGAGCAGTGACTTTGGG 405

Qy    301 AAATTTTTGTGTAAGGCTGTCCATATCATCGAACCTCAACCTCTACAGCAGCGGTTCTC 360
            ||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    406 AACTTCCTATGCAAGGCAGTCATGTCATCATCAACAGTCAACCTCTCTACAGCAGCGGTTCTC 465

Qy    361 ATCCTGGCCTTCATCAGCCTGACCGTGACCGTTCCACCGCCACCACCAACAGTCAA 420

Db    466 ATCCTGGCCTTCATCAGTCTGGACCGTGACCGCCATCGGCCTGCCACCGCCACCAACAGTCAG 525

Qy    421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGGTCTGAGATCCCAGCCCTC 480
            ||||||||||||||||||||||||||||||||||||
```

```
Qy    541 GATGACAGGTACATCTGTGACCGCCTTACCCCGATAGCCTGTGGATGGTGTGTTCAA 600
            ||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    631 GATGACAGATATCTGTGACCGCCTTCTCATCCTGCCCGGCATCGTCATCCTCGTGTTACTGC 690

Qy    601 TTCCAGCATAAATGGTGGTCTGCTTCATCCTGTATCCTCCTGTTACTGC 660
            ||||||||||||||||||| ||||||||||||||||||||||||
Db    691 TTTCAGCACATCATGGTTGGCCTTATCCTGCTGGTATTGTCATCCTGTCCTGCTATTGC 750

Qy    661 ATCATCATCTCTAAGCTGTCACACTGCAACAGAGCGCACACTCAAAGGCCCTCAAGACG 720
            ||||||||||||||||||| ||||||||||||||||||||||||
Db    751 ATTATCATCTCCCAAGCTGTCACACTGCAACAGAGCCAAGGGCCAAGACC 810

Qy    721 ACAGTCATCCTCATCCTCAGTTTCTTTGCCTGGCTGGCTGCCATATTATGTGGGGATCAGC 780
            ||||||||||||||||||| |||||||||||||||||||||||||
Db    811 ACAGTCATCCTCATCCTCGGCTTTCGCCTGTCGCCTGTTCCGTGGCTCCTTACTACACATTGGGATCAGC 870

Qy    781 ATCGACTCCTTCATCCTTTTGGAGTCATCAAGAGCAAGGATGTGACTTCTGGACGCATTGTG 840
            ||||||||||||||||||| ||||||||||||||||||||||||||
Db    871 ATCGACTCCCTTCATCCTCCTGGAAATCATCAAGCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 930

Qy    841 CACAAGTGGATCTCCATCACAGAGAGCCCTCGCCTTCTTCCACTGTGCCTGACCCCATC 900
            ||||||||||||||||||| |||||||||||||||||||||||||||
Db    931 CACAAGTGGATTTCCATCACCGAGAGCCCTAGCTTCTTCCACTGTTGTGCTGACCCCATC 990

Qy    901 CTCTATGCCTTCCTGGGGCCAAGTCAAAGCTCTGCCCAGCATGCACTCAACTCCATG 960
            ||||||||||||||||||| |||||||||||||||||||||
Db    991 CTCTATGCTTTCCTGGGAGCCCAAATTTAAAACCTCCGCCCAGCACGCACTCACCTCTGTG 1050

Qy    961 AGCAGAGGCTCCAGCCTCAAGATCCTTCTCAAAGGAAAGCAGGGGGGTGGACACTCTTCCGTC 1020
            ||||||||||||||||||| |||||||||||||||||||||||||
Db   1051 AGCAGAGGGTCCAGCCTCCTCAAGATCCTCTCCAAAGGAAAGCGAGTGCAGAGTGGAACTCTGTT 1110

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCAGCTAACCCTTATGCAAAGACTTATAT 1080
            ||||||||||||||||||| |||||||||||||||||
Db   1111 TCCACTGAGTCTGAGCTCTTCAAGTTTTCACTCCCAGCTAA-------CACAGATGTAAAA 1162

Qy   1081 AATATATATATATATGATAAATAACTTTTTTTTTTTATGTTACACATTTCCAGATATAAAAG 1140
            ||||||||||||||||||| |||||||||||||
Db   1163 GACTTTTTTTTATACGATAAATAACTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1222

Qy   1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTAATTGACTGTTGGGGAGTTTATGTTCC 1200
            ||||||||||||||||||| |||||||||
Db   1223 ACTGACCAAATATTGTACAGTTTTTATTGCTTGTGTTGGATTTTTGT------CTTGTGTTTC 1276

Qy   1201 TCTAGTTTTTGTGAGGTTTGACTAAATTTATATAAAATATTGTTTTTGTTTGTTTCATGT 1260
            ||||||||||||||||||| |||||||||||||||||
Db   1277 TTTAGTTTTTGTGAAGTTTAATTGACTTATTTATTTATA---TAAATTTTTTTTTGTTTCATAT 1332

Qy   1261 GAATGAGGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
            ||||||||||||||||||| ||||||||||||
Db   1333 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTGTGTGTATGCTCGTGGTAG 1392

Qy   1321 GACTGTAGAACTGTGAGAGGAAGAAACTGAACATTCCAGAAATGTGTGTAAATTGAAATAAA 1380
            ||||||||||||||||||| |||||||
Db   1393 GACTGTAGAA-------AAGGGAACTGAACATTCCAGAGGCGGTGTGAGTGAGTAGTGAATCACGTAA 1444

Qy   1381 GCTAGCCGTGATCCTGCAGCCTGTTGCTGCATA 1411
            ||||||||||||||||||| ||||||
Db   1445 GCTAGAAAATGATCCCCAGCTGTTTATGCATA 1475
```

```
RESULT 6
US-08-701-265-3
; Sequence 3, Application US/08701265
; Patent No. 5776457
; GENERAL INFORMATION:
    APPLICANT:  Chuntharapai, Anan
    APPLICANT:  Lee, James
    APPLICANT:  Hebert, Caroline
    APPLICANT:  Jin Kim, K.
  TITLE OF INVENTION: Antibodies to Human PF4A Receptors
  NUMBER OF SEQUENCES:  6
  CORRESPONDENCE ADDRESS:
```

```
APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.
TITLE OF INVENTION: Antibodies to Human PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Genentech, Inc.
  STREET: 460 Point San Bruno Blvd
  CITY: South San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94080
COMPUTER READABLE FORM:
  MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/284,586
  FILING DATE:
  CLASSIFICATION: 424
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/076,093A
  FILING DATE: 11-Jun-1993
  APPLICATION NUMBER: 07/810782
  FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/677211
  FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Love, Richard B
  REGISTRATION NUMBER: 34,659
  REFERENCE/DOCKET NUMBER: 706P2
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415/225-5530
  TELEFAX: 415/952-9881
  TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
  LENGTH: 1737 nucleotides
  TYPE: Nucleic Acid
  STRANDEDNESS: Single
  TOPOLOGY: Linear
US-08-284-586-3

Query Match          55.1%;     Score 933.4;     DB 2;     Length 1737;
Best Local Similarity 81.8%;     Pred. No. 1e-210;
Matches 1154;   Conservative 0;   Mismatches 216;   Indels 41;   Gaps 5;
```

```
Qy    1 ATATACACTTCGATAACTACTCTGAAGAAGTGGGGTCTGGAGCATGACTCCAACAAG    60
         ||||||||||| |||||||| ||  || ||||| ||| || ||||||| ||||| ||||
Db  106 ATATACACTTCAGATAACTACACACCGAGGAAATGGCTCAGGGACTATGACTCCATGAAG  165

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCATTCCTGCCCACCATCTAC               120
         ||||||||  |||| |||||||||  | || ||||||||||||||||
Db  166 GAACCCTGTTCCGTGGAGGAAATGCTAATTCATAAAATCTTCCTGCCCACCATCTAC     225

Qy  121 TTCATCATCTTCTTGTACTGGCATAGTCCGGCAATGATTGGTGATCCTGTCATGGGTTAC  180
         | |||||||||||| ||||||||| ||  ||||||||||| ||||| ||||||||||||||
Db  226 TCCATCATCTTCTTAACTGGCAATTGTGGGCCAATGGTGTCATCCTGTCGTCATGGGTTAC  285

Qy  181 CAGAAGAAGCTAAGGAGCTGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC   240
         |||||||| || || |||||||||||| |||| ||||| |||||||||||||| |||||||
Db  286 CAGAAGAAACTGAGAAGCATGACGACAGTAGCAGGCTCACACCTGTCAGTGGCCGACCTC  345

Qy  241 CTCTTTGTCATCACACTCCCCTTCCCTTCGGGCAGTGATGCCATGGCCGTGGTACTTTGG  300
         ||||||||||||||| || ||||||||||||||||||||||| ||||| |||||||||||
Db  346 CTCTTTGTCATCACGCTTCCCTTCCCTTCGGGCAGTGATGCCGTGGCAAACTGGTACTTTGGG  405

Qy  301 AAATTTTGTGTAAGGCTGTCCATATCATCTCACAGTCATTCTACACGTCAACCTTCACACT  360
         || || | ||||||||||||||| |||||||||||| |||||||| |||||| |||||
Db  406 AACTTCCTATGCAAGGCAGTCCATGTCAATCATCTCACACAGTCAACCCTCACAGCAGTGTCCTC  465
```

```
Db  526 AGGCCAAGGAAGCTGTTGGCTGAAAAGTGGTCTATGTTGGCGTCTGGATCCCTGCCCTC   585
Qy  481 CTCCTGACTATACCTGACTTCATCTTTGCCGACTTCATCTTTGCCAACG------       540
         || |||||||||  ||||||||||||||| ||||||||||||||||||||
Db  586 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG----------TCAGTGAGGCA       630

Qy  541 GATGACAGGTACATCTGTGACCGGCCTTTACCCCGATAGCCTGTGGATGGTGGTTGTTCAA  600
         |||||||| |||||| ||||||||||||  ||||||| | ||||| |||||||||| ||| |
Db  631 GATGACAGATATACCGTGACCGGCCTTCTACCCCAATGACTTGTGGTGTTGGTGTTCCAG   690

Qy  601 TTCCAGCATATAAATGGTGGGTCTCATCCTGTCCCGGCATCGTCATCCTCTGTACTGC     660
         || |||||||| | | || | || ||||| |||| |||||||||||||| || || |||
Db  691 TTTCAGCACATCATGGTTGGGCCTTATCCGCTGGTATTGTCATCCTGTCCTGCTATTGC   750

Qy  661 ATCACATCCTAAGCTGTCACACTCCAAGGGCCACCAGAAGGCCAAGGCCCTCAAGACG    720
         || |||||||| ||||||||||||||||||||||||| ||||||||||||||||||||
Db  751 ATTATCATCTTCACAGCTGTCACACTCCAAGGGCCACCACCGAAGGGCCAAGGCCCTCAAGACC  810

Qy  721 ACAGTCATCCTCATCCTAGCTTTCTTGCCTGCTGGCTGCCATATTATGTGGGATCAGC    780
         |||||||||||||||| | ||||||||||||||||||||| | |||| |||||||||||
Db  811 ACAGTCATCCTCATCCTCGGCTTTCTTGCCTGCTGGCTGCCTTACTACATTGGGATCAGC   870

Qy  781 ATCGACTCCTTCATCCTTTGGGAGTGCATCAAGCAAGGATGTGACTTCGAGAGGCATTGTG  840
         |||||||||||||||||| | | | || ||||||||| ||||| ||| ||||| |||||
Db  871 ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGCTGTGAGTTGGAGAACACTGTG   930

Qy  841 CACAAGTGGATCTCCATCACAGAGGCCCTTCTTCCACTGTTGCCTGAACCCCATC       900
         |||||||||| ||||||| |||||||||  ||||||||||| || ||||||||||
Db  931 CACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTCTTCCACTGTTGTCCAGAACCCCATC  990

Qy  901 CTCTATGCCTTCCTCGGGGCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG    960
         ||||||||  ||||| || ||||  |||||  ||||||||||| |||||| || ||||||
Db  991 CTCTATGCTTTCCTTGGAGCCAAATTTAAAACCTCTGCCCAGCACGCACTCACCTCTGTG  1050

Qy  961 AGCAGAGCCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC  1020
         |||||||   ||||||||||||||| |||||||||||||| | ||||| ||| | | |||
Db 1051 AGCAGAGGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGAGCATTCATCTGTT  1110

Qy 1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT   1080
         |||||  |||| ||  | || |||||||||||||||||       |||| | |||  ||
Db 1111 TCCACTGAGTCTGAGTTTCAAGTTTCACTCCAGCTAA-------CACAGATGTAAA     1162

Qy 1081 AATATATATATATATACGATAAATAACTTTTTTTTTATGTTACACATTTCCAGATATAAGAG  1140
         ||||||||||||||| ||||| || |||||||||||  |||||| |||||||| |||| ||
Db 1163 GACTTTTTTTTTTATCAGTTTTTTTTTTTTTTTAATTGACTGTGTGGGAGTTTATGTTCC   1222

Qy 1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTATTGCTTGTTGGATTTTTGT------CTTGTGTTTC 1200
         |||||||| | |||||||||||||  | || |||||| ||||| |||||||||      |||||||||
Db 1223 ACTGACCAATATTGTACAGTTTTTATTGGTAAATTGATTTTTGT------CTTGTGTTTC       1276

Qy 1201 TCTAGTTTTTGTGAGGTTTGACTTAATTTATATAATATATTGTTTTTTGTTTGTTTCATGT  1260
         | |||||||| |||||| || |||||||||||     ||| ||| || || |||||||| |
Db 1277 TTTAGTTTTTTGTGAGAGTTTAATTGACTTAATTTATA----TAAATTTTTTTGTTTCATAT  1332

Qy 1261 GAATGAGCGTAGGACGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG     1320
         | |||| ||  ||||| |||||||||||||||||||||| ||| |||| ||| |||||
Db 1333 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGTGTCTATGTCTCGTGGTAG   1392

Qy 1321 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA   1380
         |||||||||       | |||| |||||||||||||||| | || ||||||||||| ||||
Db 1393 GACTGTAGAA--------AAGGGAACTGAACATTCCAGACGCGTGAGTGAGTCGTGAATCACGTAA  1444

Qy 1381 GCTAGCCGTGATCCTGTTGCTGCATA                                   1411
         ||||| | ||||||| |||||| ||||
Db 1445 GCTAGAAATGATCCCCAGCGTGTTTATGCATA                             1475
```

RESULT 7
US-08-284-586-3
; Sequence 3, Application US/08284586
; Patent No. 5840856
; GENERAL INFORMATION:
;   APPLICANT: Chuntharapai, Anan

RESULT 8
US-08-805-478-3
; Sequence 3, Application US/08805478
; Patent No. 5874543
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,478
; FILING DATE: 25-Feb-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1C1
; TELECOMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-805-478-3

Query Match   55.1%;   Best Local Similarity   81.8%;   Score 933.4;   DB 2;   Length 1737;
Matches 1154;   Conservative   0;   Mismatches 216;   Indels   41;   Gaps   5;

```
Qy     1 AATATACACTTCTGATAACTACTCTGAGAAGAAGTGGGTCTGAAGAAGAAGTGATGACTCCAACAAG    60
         ||||||||||| |||||||||| || || || ||| |||| || || |  |||||||||||| ||| |
Db   106 AATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG       165

Qy    61 GAACCCTGCTTCCGGATGAAAACGTCATTTCAATGGATCTTCCTGCCACCATCTAC              120
         |||||||| ||||| ||| || ||  ||  |||||  ||||||||| |||||||||
Db   166 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTAC          225

Qy   121 TTCATCATCTTCTTGACTGGTCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC         180
         | |||||||||||| |||||||||| | ||||||||||||||| |||||||||||||||||
Db   226 TCCATCATCTTCTTAACTGGTCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC        285

Qy   181 CAGAAGAAGCTAAGGAGGAGCATGACGGCAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC        240
         ||||||||| || ||| ||||||||| || ||||||| ||||||||||||||||||| ||||||
Db   286 CAGAAGAAACTGAGAAGGCATGACGGACAAGTACAGCGGCTGCACCTGTCAGTGGCCGACCTC     345

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGTGGTACTTTGGG               300
```

```
Qy   361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAA          420
         |||||||||||||||||| || ||| | |||||||||| ||||||||||||||||||||||
Db   466 ATCCTGGCCTTCATCAGTCAGTCTGGCCGCTGGCCATCGTGTGGTGTTGTTCCAACAGTCAG       525

Qy   421 AGGCCAAGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCCAGCCCTC         480
         |||||||||||||||||||||||||||||||||||| ||| | |||||| ||||||||||
Db   526 AGGCCAAGGAAGCTGGTTGGCTGAAAAGGTCGTCTATGTGTGGCGTCTCGGATCCCTGCCCTC    585

Qy   481 CTCCTGACTATACCCGACTTCATCTTGCCGACGTCAGCCAGGGGGACACTCAGTCAGGGG        540
         ||  |||    |||  |||  | |||| |||        ||
Db   586 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG-------TCAGTGAGGCA             630

Qy   541 GATGACAGTGACATCTGTGACCGCCTTTACCCCGATAGCCTGTGGATGGTGTGTTTCAA         600
         |||||||| |||||| ||||| ||| |  || || || ||  ||||||| |||||||||
Db   631 GATGACAGATATATCTGTGACCGCGCTTCTACCCCAATTGCTGTGGTGTGTTTGTTCCAG      690

Qy   601 TTCCAGCATATAATGGTGGTCTCATCATCCTGCTCATCCTCCTGTTACTGC                660
         |||||||| |||  |||||||||||| |||||||  |||| |||||| ||||
Db   691 TTTCAGCACATCATCGGTTGGCCTTATCATCGTCATCCTCGTCCTCGTTACTGC            750

Qy   661 ATCATCATCTCTAAGCTGTCACACTCCCAGAGGCCACCAAGGGCCCTCAAGACG             720
         ||| |||||||| || |||||  |||||| || || |||| ||| |||||| ||
Db   751 ATTATCATCTCTCAAGCTGTCACACTCCCAAGGGCCAAGGGCCAAGGGCCCCAAGACC        810

Qy   721 ACAGTCATCCTCCATCACCTAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGC     780
         |||||||||||| ||||| ||||| ||||| ||| || || ||||| || |||| |||||||||
Db   811 ACAGTCATCCTCTCATCACCTGGCTTTCTTTGCCTGCTGTGTGCCTTACAATTGGGCTGCTCAGC 870

Qy   781 ATGGACTCCTTCATCCTTTTGGGAGTCATCAAGAGACAAGGATGTGACTTCGAGAGCATTGTG    840
         ||| || |||||||||| |||| |||||| |||| || |||| ||||| |||||||| |||||
Db   871 ATCGAGTCCTTCATCCTCCTGGAAATCATCAAGCAGCAAGGCTGTGAGTTTGAGAGTGTTGTG    930

Qy   841 CACAAGTGGATCTCCATCACAGAGGCCCTCGCTTCTTCCACTGTTGCCTGAACCCCATC       900
         |||||| ||| |||||| |||| ||| || || ||||||||| |||| ||||||||||
Db   931 CACAAGTGGATCTCCATCACACCGCGAGGCCCTAGCTTCTTCCACTGTTGTGAACCCCATC       990

Qy   901 CTCTATGCCTTCCTCGGGGGCAAGTTCAGAAAGAGCTCACACGCACTCAACTCCATG        960
         || |||||||| || || |||| |||||||||| ||  || ||  ||||| ||||||
Db   991 CTCTATGCCTTCCTCGGAGCAGCCAAATTTAAAACCTCGGCACAGCACGCACTCCTGTG        1050

Qy   961 AGCAGAGGCTCCAGCCTCAAGATCTCTTCCAAAGGAAAGGCGGGGTGGCCACTCTTCCGTC      1020
         ||||||||||||||||| ||||| || || |||||||||| ||||| || || ||||| ||
Db  1051 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGCGGTGTGTGTGGACTCTGTT    1110

Qy  1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCAGCTAACCCTTATGCAAAGACTTATAT       1080
         ||||| || |||| ||| || ||||||||  |   || ||||||||
Db  1111 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCAAGT------CACAGATGTAAAA            1162

Qy  1081 AATATATATATATATGATAAATAACTTTTTTTTTTTTTTTTAAGTTACACATTTCAGATATAGAG    1140
         |||||||||| || ||||||| ||||| ||||||||||||||| |||||||| ||||
Db  1163 GACTTTTTTTTTTATACGATAATAACTTTTTTTTTTTTTTTTAAGTTACACATTTCAGATATAAA  1222

Qy  1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTTTTTAATTGACTGTGTGGGAGTTTATGTTCC    1200
         |||||| ||| |||| ||||||||||||||||||  |  ||||| |||| |   |||| |||
Db  1223 ACTGACCAATATTGTACAGTTTTTTTTTTATTGCTTGTCGTGTGGATTTTTTGT------CTTGTGTTTC  1276

Qy  1201 TCTAGTTTTTGTGAGTTTGACTAATTATATAAATATGTTTTGTTTTTTGTTTCATGT         1260
         | ||||||||||||||| |||||||||||||  ||     ||||||||||| ||||
Db  1277 TTTAGTTTTTGTGAGTTTGAAGTTTAATTGACTTATTTATA------TAAATTTTTTTTGTTTCATAT   1332

Qy  1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTAG        1320
         |||| | || ||||||||| || ||| ||||||| ||| |||||||| ||  || |||
Db  1333 TGATGTGTGTTCCGTCAGGCAGGACCTGTGGGCCAAGTTCTTAGTTGCTGTCTATGCTCGTGGTAG   1392

Qy  1321 GACTGTAGAACTGTGTAGAGGAAGAAGAAACATCCAGAAATGTGTGGTAAATTGAATAAA      1380
         |||||| ||    || |||||| ||||| |||||| |   | |||| |||||  ||||
Db  1393 GACTGTAGA------AAGGGAACTGTAGAAGCATTCAGAGCCGTGTAGTGGAATCACGTAAA      1444

Qy  1381 GCTAGCCGTGGATCCTCAGTCGTTGCTGCATA             1411
         |||||| | ||  | || |||| |||||||||
Db  1445 GCTAGAAATGATCCCCAGCTGTTTATGCATA                1475
```

Db   1393 GACTGTAGAA--------AAGGGAACTGAACATTCCAGAGCGTGTAGTGAATCACGTAAA 1444

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
            ||||||||||||||||||||||||||||||||
Db   1445 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1475

RESULT 9
US-08-802-627A-3
; Sequence-3, Application US/08802627A
; Patent No. 5892017
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,627A
; FILING DATE: 19-Feb-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-802-627A-3

Query Match   55.1%; Score 933.4; DB 2; Length 1737;
Best Local Similarity 81.8%; Pred. No. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAGAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
          ||||||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||
Db   106 ATATACACTTCAGATAACTACCACCGAGGAAATGGGGCTCAGGGGACTATGACTCCATGAAG 165

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCCTGCCCACCATCTAC 120
          ||||| |||||||||||||||||||| || |||||||||||||||||||
Db   166 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCCACCATCTAC 225

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   226 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 285

Db   1333 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTCTTCTTAGTTGCTGTATGTCTCCTGGGTAG 1392
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321 GACTCTAGAACTGTGAGAGGAAGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
          ||||||||||| ||||||||||||||||||||||
Db   1393 GACTGTAGAA------AAGGGAACTGAACATTCCAGCGTGTAGTGAATCACGTAA 1444

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
          |||| |||||||||||||||||||||||||
Db   1445 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1475

RESULT 10
US-08-801-238-3
; Sequence 3, Application US/08801238
; Patent No. 5919896
; GENERAL INFORMATION:
;   APPLICANT: Lee, James
;   APPLICANT: Wood, William I.
;   TITLE OF INVENTION: PF4A RECEPTOR
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/801,238
;   FILING DATE: 19-Feb-1997
;   CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/284586
;   FILING DATE: 10-AUG-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/076093
;   FILING DATE: 11-JUN-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/810782
;   FILING DATE: 19-DEC-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B.
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: P0706P2P1D1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1737 base pairs
;   TYPE: Nucleic Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Linear
;
US-08-801-238-3

Query Match   55.1%; Score 933.4; DB 2; Length 1737;
Best Local Similarity   81.8%; Pred. No. 1e-210;
Matches 1154; Conservative   0; Mismatches 216; Indels 41; Gaps   5;

Qy   1 ATATACACTTCTGATAACTACTCTGGAGAAGTGGGGTCTGGAGACTATGACTCCAACACAAG 60
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   106 ATATACACTTCAGATAACTAACTACACCAGGACTCACCAGGGACTATGACTCCATGAAG 165

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 120
          ||||| |||| ||||||||||| |||||||||||||||||||||||||||||||||||||
Db   166 GAACCCTGTTTCCGTGAAGAAATGCTAATTTCAATAAAAATCTTCCTGCCCACCATCTAC 225

Qy   181 CAGAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
          ||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||
Db   286 CAGAGAAACTGAGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCCGACCTC 345

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGATGCCATGGCCATGTACTTTGGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 CTCTTTGTCATCACGTTCCCTTCTGGGCAGTGCAAACTGTACTTTGGG 405

Qy   301 AAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGGTTCTC 360
          || ||||||||||| |||||||||||||||||||||||||||||||||||||
Db   406 AACTTCCTATGCGAAGGCAGTCCATGTCATCTACACAGTCAAACCTCTACAGCAGTGTCCTC 465

Qy   361 ATCCTGCCTTCATCAGCTGACCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAA 420
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   466 ATCCTGCCTTCATCAGCTGACCTCGGACCGTCGGCCATGGCCACCCAGCCACCAACAGCTCAG 525

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
          ||||||||||||||||||||| |||||||| |||||||||||||||||||||||||||
Db   526 AGGCCAAGGAAGCTGTTGGCTGCAAAGGTGGTCTATGTTGGCGTCTCGGATCCCTGCCCTC 585

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGTCAGCCAGGGGACATCAGTCAGGGG 540
          |||||||||||||||||||||||||||||||||| ||||||||||||
Db   586 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG-----------TCAGTGAGGCA 630

Qy   541 GATGACAGGTACATCTGTGACCGCCTTTACCCGGATTAGCTGTGGATGGTCGTGTTTCAA 600
          ||||||||||| |||| ||| ||||| || ||||||| || |||| || ||||||
Db   631 GATGACAGCAGATATCTGTGACCGGTACTTTACCCCCAATTGCGTGGTTGGTTGTTCCAG 690

Qy   601 TTCCAGCATATAAATGGTGGGTCTCATCCTCGCCGCCGGATCGTCATCCTCTCCTGTTACTGC 660
          |||| ||| | | | |||| |||||||||||| |||||||||||||||||||||
Db   691 TTTCAGCACATCATGGTTGGCCTTATCCTGTCTGCCTTGTCATCCTGCTATTGC 750

Qy   661 ATCATCATCTCTAAGGCTGTCACACTCCAAGCTGCCACCAGAGCGCCACGAGGCCCTCAAGACG 720
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   751 ATTATCATCTCCAAGTGTCACACTCCAAGGGCCCACCAGAGAGCGCAAGGGCCCTCAAGACC 810

Qy   721 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   811 ACAGTCATCCTCATCCTTGCCTTCTTCGCCTGTTGGCTGCCTTACTACATTGGGATCAGC 870

Qy   781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
          |||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db   871 ATCGACTCCTTCATCCTCTGGAAATCATCAAGCAAGGTGTGTGAGTTTGAACACTGTG 930

Qy   841 CACAAGTGGATCTCCATCACAGAGAGGCCCTGCCTTCTTGTTGCCTGAACCCCATC 900
          ||||||||||||| |||||||||||||||| |||||||||||||||||||||||
Db   931 CACAAGTGGATTTCCATCACCGAGGCGCCCTGCTTCTTGTTGCTGCTGTTGTCTGAACCCCATC 990

Qy   901 CTCTATGCCTTCTCCGGGGCCAAGTTCAAAAGTCTGCCCAGCAGCTCTGCCAGCACTCCAAGTTCCCATG 960
          |||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||
Db   991 CTCTATGCTTTCTTCCTTGGAGCCAAATTTAAAACCTCGCCCCAGCAGCACGCACCTCTGTG 1050

Qy   961 AGCAGAGGCTCCAGCCTCAAGATACCTTTCCAAAGGAAGCGGGGTGACAACTCTTCCGTC 1020
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   1051 AGCAGAGGGTCCAGCCTCAAGATCCTCCAAAGGAAGCGAGGTGGCAGATGTGACACTCGTT 1110

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTTCACTCCAGAGTCAACCTTATGCGAAAGACTTATAT 1080
          ||||||||||||||||| |||||||| |||||| ||||||||||||||||||||
Db   1111 TCCACTGAGTCTGAGTCTTCAAGTTTTCACTCCAGCTAA------CACAGATGTAAAA 1162

Qy   1081 AATATATATATATATGATAAAGAACTTTTTTATTGTTTAGAGTTACACATTTTCCAGATATAAGAG 1140
          |||| ||||||||||||||||||| ||||||||||||||||||||||||
Db   1163 GACTTT-TTTTATTCGATAAATAACTTTTTTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1222

Qy   1141 ACTGACCAGCTCTGTGACAGTTTTTTTTTTTTTTTGTTGT------CTTGTGTTTC 1276
          ||||||||||||| ||| | |||||||||| ||||||| |||||||||
Db   1223 ACTGACCAATATTGTACAGTTTTTGCTTGTTGGATTTTGT------CTTGTGTTTC 1276

Qy   1201 TCTTAGTTTTGTGAGGTTGACTAAATTAATTATATATAAATATTGTTTTTTTGTTTTTCATGT 1260
          ||| ||||| |||||||| |||||||||||||||||||| ||||||||||||||
Db   1277 TTTTAGTTTTTTGTGAAGTTAATTGACTTTATTTATTTATA----TAAATTTTTTTTGTTTCATAT 1332

Qy   1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTTATCTGTGTGTAG 1320
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1261 GAACCCTGTTTCCGTGAAGAAATGCTAATTTCAATAAAAATCTTCCTGCCCACCATCTAC 225

```
Qy  1201 TCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAATATTGTTTTTTTGTTGTTTCATGT 1260
              ||||||||||||| ||||  ||||| | ||||||||| |   |||||||||||||||||
Db  1277 TTTAGTTTTTGTGAAGTTTAATTGACTTATTTTATA----TAAATTTTTTTTGTTTCATAT 1332

Qy  1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
              |||| | ||||||||||||||||||||||||||||||||| |||||| |||||| |||
Db  1333 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTGTGTCTCGTGGTAG 1392

Qy  1321 GACTGTAGAACTGTAGAGGAGAAACTGAACATTCCAGAATGTGGTAAATTGAATAAA 1380
              |||||||||              |||||||||||||||  ||||| |||  | ||||
Db  1393 GACTGTAGAA-------AAGGGACTGAACATTCCAGAGCGTGTAGTGAATCACGTAAA 1444

Qy  1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
              ||||| | ||||| |||||||| | |||||
Db  1445 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1475
```

RESULT 11
US-08-801-228-3
; Sequence 3, Application US/08801228
; Patent No. 5922541
; GENERAL INFORMATION:
;   APPLICANT: Lee, James
;   APPLICANT: Wood, William I.
;   TITLE OF INVENTION: METHODS FOR DETECTION AND AMPLIFICATION OF
;   TITLE OF INVENTION: PF4A RECEPTOR NUCLEIC ACID
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Genentech, Inc.
;   STREET: 460 Point San Bruno Blvd
;   CITY: South San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94080
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/801,228
;   FILING DATE: 19-Feb-1997
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/284586
;   FILING DATE: 10-AUG-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/076093
;   FILING DATE: 11-JUN-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/810782
;   FILING DATE: 19-DEC-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Love, Richard B.
;   REGISTRATION NUMBER: 34,659
;   REFERENCE/DOCKET NUMBER: P0706P2P1D3
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-5530
;   TELEFAX: 415/952-9881
;   TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1737 base pairs
;   TYPE: Nucleic Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Linear
US-08-801-228-3

Query Match   55.1%; Score 933.4; DB 2; Length 1737;
Best Local Similarity 81.8%; Pred. No. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

```
Qy     1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAG 60
```

```
Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGGGTTAC 180
              ||||||||||||| |||||||| | ||||||| |||||| |||||||||||||||||
Db   226 TCCATCATCTTCTTAACTGGCATTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 285

Qy   181 CAGAAGAAGCTAAGGAGCATGACGGACAAGCTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
              ||||||||| || || |||||||||||||||| || | | |||||||||||| ||||||
Db   286 CAGAAGAAACTGAGAAGCATGACGGACAAGCTGCACCTGCACCTGTCAGTGGCCGACCTC 345

Qy   241 CTCTTTGTCATCACACTCCCCTTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 300
              ||||| || |||| ||||| |||||||| |||| |||| || ||||| ||||||||||
Db   346 CTCTTTGTCATCACGCTTCCCTTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG 405

Qy   301 AAATTTTTGTGTAAGGCTGTCCATATCATCCTACACGTCTCAACCTCTACAGCAGCGGTTCTC 360
              || |||| |||||||||||||||| |||||| ||||||||||||||||||||||||||| |||
Db   406 AACTTCCTATGCAAGGCAGTCCATGTCATCACAGTCTCAACCTCTACAGCAGTGTCCTC 465

Qy   361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTGGCCATTGTCCACGCCACCAACAGTCAA 420
              |||||||||||||||||| ||||||| |||||||||||||||||| ||||||||||||||
Db   466 ATCCTGGCCTTCATCAGCCTGGACCGCTACCTGGCCATTGTCCACGCCACCAACAGTCAG 525

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCCAGCCCTC 480
              ||||||||||||||||||||||||||||| ||| |||||||| ||||||||||||||||||
Db   526 AGGCCAAGGAAGCTGCTGGCTGAAAAGGTGTATGTGGGTGTCTGGATCCCTGCCCTC 585

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCAGCGGGGACAGTGTCAGTGAGGGG 540
              ||| |||||| |||||||||||||||||||| ||||||||||||
Db   586 CTGCTGACTATTCCCGACTTTCATCTTTGCCAACG----------TCAGTGAGGGCA 630

Qy   541 GATGACAGGTACATCTGTGACCGCCTTACCCCCGATAGCCTGTGATGGTGGTGTTTCAA 600
              |||||||| ||||| ||||||| || || || || || ||||||||||| || |||||
Db   631 GATGACAGATATCGTGACCGCTTCTACCCCAATGACTGTGGTGGTGGTGTTGTTCCAG 690

Qy   601 TTCCAGCATATAAATGGTGGGTCATCCTCTCCTCCTGTTACTGC 660
              ||||||| | || ||||| ||||| |||||||||||||| ||||
Db   691 TTTCAGCACATCATGGCCTTATCCTGTATTCTCCTGCTATTGC 750

Qy   661 ATCATCATCTCTAAGGCTGTCACTGTCCAGAAGGCGCCTCAAGACG 720
              ||| |||||| ||||| |||||||| ||||||||||| ||||||||
Db   751 ATTATCATCTTCAAGTGTCACTGTTCAAGAAGGCGCCCAAGACC 810

Qy   721 ACAGTCATCCTCCAGCTTTCTTGCCTGCTGGCCATTATGTGGGGATCAGC 780
              ||||||||||||||| ||||||||| |||||||||||||||||||||||||
Db   811 ACAGTCATCCTCCAGCTTTCTTGCCTGCTGGCCATTATGTGGGGATCAGC 870

Qy   781 ATCGACTCCTTCATCCTTTGGGAGTGATCAAGCAAGGATGGACTTCGAGAGCATTGTG 840
              ||| ||||||| |||||||||||| |||||||||||| || ||||| |||| ||| |||
Db   871 ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 930

Qy   841 CACAAGTGGATCTCCATCACAGAGAGCCCCGCCTTCTTCCCACTGTTGCCTGAACCCCATC 900
              |||||||||||||||||| ||||||||| ||||||||||||||||| ||||| ||||||||
Db   931 CACAAGTGGATTCCATCACACCGAGGCCTAGCTTCTTCTTCCACTGTGTCTCGAACCCCATC 990

Qy   901 CTCTATGCCTTCCTGGGGCCAAGTTCAAAAGCCTGCCCAGCATGCACTCAACTCCATG 960
              ||| | ||||| |||||||||||||||||||| ||||| |||| |||||||||||| ||
Db   991 CTCTATGCTTTCCTGGGAGCCAAATTTAAAACCTGCCCAGCACGCACCTGGTGTGTG 1050

Qy   961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAAGCCGGGGGTGGACACTCTTCCGTC 1020
              |||||||| ||||||| |||||| || ||||| |||||||| || |||||||||||| ||||
Db  1051 AGCAGAGGGTCCAGCCTCAAGATCCTCCAAAGGAAAAGCGAGGTGGACATTCATCTGTT 1110

Qy  1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCCAGCTAACCCTATGCAAAGAGACTTATAT 1080
              |||| ||||||| |||||| || |||||||| |              || ||||| |||| 
Db  1111 TCCACTGAGTCTGAGTCTTCAAGTTTTCACTCCCAGCTAA-------CACAGATGTAAAA 1162

Qy  1081 AATATATATATATATGATGATAAAGAAGAACTTTTTTTTTTTTATGTTACACAATTTCCAGATATAAGAG 1140
              |||||||||||||||||||| |||||| ||||||||||||||||||| |||| ||||||||||||||||||
Db  1163 GACTTTTTTTTTATACGATAAATAACTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1222

Qy  1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTTAATTGACTGTGGGAGTTTATGTTCC 1200
              |||||||| ||| || |||||||||||||||| ||||||||| ||||||||||
Db  1223 ACTGACCAAATTGTACAGTTTTTTATTGCTTGTGGATTTTTGT------CTTGTGTTTC 1276
```

Db 1163 GACTTTTTTTATACGATAAATAACTTTTTTTTTTTTTTTTTTTTAAGTTACACATTTTTCAGATATAAAAG 1222
Qy 1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTTTTTTAATTGACTGTTGTGGAGTTTTATGTTCC 1200
Db 1223 ACTGACCAATATGTACAGTTTTTATTGCTTGTTGGATTTTTGT------CTTGTGTTTC 1276

Qy 1201 TCTAGTTTTTGTGAGGTTGACTTGACTTAATTATATAAATATTGTTTTTGTTGTTTCATGT 1260
Db 1277 TTTAGTTTTTGTGAAGTTAATTGACTTATTTATA----TAAATTTTTTTGTTTCATAT 1332

Qy 1261 GAATGAGCGCTCAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCGTCTTGTCGTGTGTAG 1320
Db 1333 TGATGTGTGTCAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTGCTGTATGTCCGTGGTAG 1392

Qy 1321 GACTGTGAGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
Db 1393 GACTGTAGAA--------AAGGGAACTGAACATTCCAGAGCCTGTAGTGAATCACGTAAA 1444

Qy 1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
Db 1445 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1475

RESULT 12
US-09-104-296-3
; Sequence 3, Application US/09104296
Patent No. 6087475
; GENERAL INFORMATION:
APPLICANT: Lee, James
APPLICANT: Wood, William I.
TITLE OF INVENTION: PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/104,296
FILING DATE: 24-June-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/701265
FILING DATE: 22-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/664228
FILING DATE: 06-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single

Db 106 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGACCATGACTCCATGAGG 165
Qy 61   GAACCCTGCTTCCGGGATGAAAAGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
Db 166 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC 225

Qy 121 TTCATCATCTTCTTGACTTCATCTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
Db 226 TCCATCATCTTCTTAACTGGCAATGGGTGGGCAATGATTGGTCATCCTGGTCATGGGTTAC 285

Qy 181 CAGAAGAAGTCAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
Db 286 CAGAAGAAACTGAGAAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC 345

Qy 241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCATGGTACTTTGGG 300
Db 346 CTCTTTGTCATCATCACACCGCTTCATGGTCATCCGTGGGCAGTTGATACTTTTGGG 405

Qy 301 AAATTTTGTGTAAGGCTGTCCATATCATCTCTACAGCAGCGTTCTC 360
Db 406 AACTTCCTATGCAAGGCAGTCATGTCATCTCACACAGTCAACTTCTACAGCAGTGTCCTC 465

Qy 361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCGCCACCTCCGCCACCAACAGTCAA 420
Db 466 ATCCTGGCCTTCATCAGTCTGGACCGGTACCTCGCCACCGCCACCAACAGTCAG 525

Qy 421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
Db 526 AGGCCAAGGAAGGAAGCTGTTGGCTGGAAAAGGTGTCTATGTGGCGTCTGGATCCCTGCCCTC 585

Qy 481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
Db 586 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG---------TCAGTGAGGCA 630

Qy 541 GATGACAGGTACATCTGACGCGCCTTACCCCGATAGCCTGTGATGGTGGTGTTTCAA 600
Db 631 GATGACAGATATATCTGTGACCGGCTTCACACCCAATGACTTGTGGGTGGTTGTGTTCCAG 690

Qy 601 TTCCAGCATATAATGGTGGGTCCTCATCCTGCCCGGCCATCGTCATCCTCTCCTGTTACTGC 660
Db 691 TTTCAGCACATCATGTTGGCCTTATCCTGCTTATCATCCTGTTCCTGCTATTGC 750

Qy 661 ATCATCATCTCTAAGCTGTCACTCCAAGAGTGTGTACACTCCAAGGCGCCTCAAGACG 720
Db 751 ATTATCATCTCCAAGCTGTCACACTCCCAAGGGCCACCAGAAGGCCAAGGCCCTCAAGACC 810

Qy 721 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGCCATATATGTGGGGATCAGC 780
Db 811 ACAGTCATCCTCATCCTCGGCTTCTTCGCCTGTTGGCTGCCTTACTACTACATTGGGATCAGC 870

Qy 781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGGCAGTGACTTGTGAGAGCATTGTG 840
Db 871 ATCGACTCCTTCATCCTCTCCTGGAAATCATCAAGCAAGGTGTGAGTTGAGAACACTGTG 930

Qy 841 CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTGCCTGAACCCCATC 900
Db 931 CACAAGTGGATTTCATCACCGAGGCCCTAGCTTTCTTCTTCCACTGTGCCTGAACCCCATC 990

Qy 901 CTCTATGCCTTCCTCGGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG 960
Db 991 CTCTATGCTTTCCTTGGAGCCAAATTTAAAACCCTGCCCAGCACACGCACTCACCTCTGTG 1050

Qy 961 AGCAGAGAGGCTTCCAGGCCTCAGAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
Db 1051 AGCAGAGGGTCCAGCCTCAGAGATCCTCCCAAAGGAAAGCGAGGTGGACATTCATCTGTT 1110

Qy 1021 TCCAACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
Db 1111 TCCACTGAGTCTGAGTCTTGAGTGTTTCACTCCAGCTAA------CACAGATGTAAAA 1162

Qy 1081 AATATATATATATATGATAAAAGAACTTTTTTTATGTTACACATTTTCCAGATATAAGAG 1140

```
;     TOPOLOGY: Linear
US-09-104-296-3

Query Match        55.1%;   Score 933.4;  DB 3;  Length 1737;
Best Local Similarity  81.8%;   Pred. No. 1e-210;
Matches 1154;  Conservative  0;  Mismatches 216;  Indels  41;  Gaps  5;

Qy      1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGGTCTGAGACTATGACTCCAACAAG 60
            |||||||||||| |||||| ||||||||||||||||||||||||||||||||||||
Db    106 ATATACACTTCAGATAACTACACCGAGGAAATGGCTCAGGGGACTATGACTCCATGAAG 165

Qy     61 GAACCCTGCTTCCGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 120
            ||||||||||||||||| |||||||| |||||||||||||||||||||||||||||
Db    166 GAACCCTGTTTCCGTGAAGAAAATGTCTAATTTCAATAAAATCTTCCTGCCACCATCTAC 225

Qy    121 TTCATCATCTTCTTGACTGCATAGTCGGCAATGGTGATCCGTGTCATCGGTTAC 180
            |||||| ||||||||||| ||||||||||||||||||||||||||||||||||||
Db    226 TCCATCATCTTCTTAACTGGCATTGGTGGCAATGGTCATCCTGTCATCGGTTAC 285

Qy    181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 CAGAAGAAACTGAGGAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC 345

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCAAACTGGTACTTTGGG 300
            ||||| |||||||||| |||||||||||||||||| ||||||||||||||||||||||||
Db    346 CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG 405

Qy    301 AAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTGTTCTC 360
            |||||||| ||||| |||||||| ||| || |||||||||||||||||||||||||||||||
Db    406 AACTTCCTATGCAAGGCAGTCGATATCATCTACACAGTCAACCTCTACAGCAGCAGTGTTCCTC 465

Qy    361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCCACGCCACCACAGTCAA 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCCACGCCACCACAGTCAG 525

Qy    421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGCGTCTGAATCCCAGCCCTC 480
            |||||||||||| || || |||||||||||| || |||| |||||||||||||||||||
Db    526 AGGCCAAGGAAGCTGTTGGCTGAAAAAGTGTCTTATGTTGCGTCTGGATCCCCGCCCTC 585

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTGCAGCCAGGGGACAATCAGTCAGGGG 540
            ||||||||||||||| || |||||||||||||||||||||||| |||
Db    586 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG----------TCAGTGAGGCA 630

Qy    541 GATGACAGGTACATCTGTGACCGCCTTACCCCGATAGCCTGTGATGGTGGTGTTTCAA 600
            ||||||||| || ||||||||||||||||| |||| |||| || ||| ||||| |||||
Db    631 GATGACAGATATATCTGTGACCGCCTTCTACACCCCAATGACTTGTGGGTGGTGTGTTCCAG 690

Qy    601 TTCCAGCATAATAATGGTGGGTCTCATCCTTCGCCGGCATCGTCATCCTCCTCGTTACTGC 660
            |||||| ||| ||||||| |||| || |||||||||||||||||||||||||||||||||||
Db    691 TTTCAGCACATCATGGTTGGGCCTTATCCTCAAGGGCCACCAGAAGCGCAAGGCGCCATTGC 750

Qy    661 ATCATCATCTCTAAGTCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGAGGCGCCTCAAGACG 720
            ||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    751 ATTATCATCTCCAAGTGTCACACTCCTCCACTGTTTCTTCTTCCACTGTTTGTCTCACCCCATC 810

Qy    721 ACAGTCATCATCCTCATCCTCGGCTTTCTTTGCCTGCTGCCATATTATGTGGGGATCAGC 780
            |||||||| |||||||||||| ||||||| |||||||||||||||||||||||||||||||
Db    811 ACAGTCATCATCCTCGGCTTTCTTTGGGAGTCAAGGATGTGACTTCGAGAGCACATTGTG 870

Qy    781 ATCGGACTTCCTTCATCCTTTGGGAGTCAAGCAAGGATGTGACTTCGAGAGCACATTGTG 840
            ||||| ||||||| ||||| || ||||||| ||||| ||||||||||||||||||||||||
Db    871 ATCGGACTTCCTTCATCCTTCCTGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 930

Qy    841 CACAAGTGGATCTCCATCACAGAGAGCCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATC 900
            |||||| |||| |||||||||||||||||||||| ||||||||||||||||||||||| ||
Db    931 CACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTTCTTCCACTGTTGTCTCACCCCATC 990

Qy    901 CTCTATGCCTTCCTCGGGGCCCAAGTTCAAAAGCTCTGCCCAGGTCATCTCGCCCAGTCA 960
            |||||||||||||| ||||| ||||||||||||||| ||||||||||| |||||||| ||
Db    991 CTCTATGCCTTCCTTGGAGCCCAAATTTAAACCTCTGCCCAGCACCACCTCCTCTGTG 1050

Qy    961 AGCAGAGGGTCCAGCCTCCAAGATCCTTTCCAAAGGAAAGCGGGGTGACACTCTTCCGTC 1020
            ||||||||||||||||||||||||||||| || |||
Db  1051 AGCAGAGGGTCCAGCCTCCAGATCCTCTCCAAAGAAAGCGGAGGTGGACATTCATCTGTT 1110

Db  1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTACCTATGCAAACCTTATGCAAAGACTTATAT 1080
            ||||||||||||||||| |||||||||||||||||| || |||||||| |-----CACAGATGTAAAA 1162

Db  1111 TCCACTGAGTCGAGTCTTCAAGTTTTCACTCCAGCTAA------CACAGATGTAAAA 1162

Qy   1081 AATATATATATATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAAAGAG 1140
            ||||||||||||||||| |||||||||||||||| |||||||||||||||||||| |||
Db   1163 GACTTTTTTTTATACGATAAATAACTTTTTTTTTTAAGTTACACATTTTTCAGATATAAAAG 1222

Qy   1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTAATTGACTGTGTGGAGTTTATGTTCC 1200
            ||||||||| || ||||||||||||||||||       |||||| ||
Db   1223 ACTGACCAAATTGTACAGTTTTTATTGGACTTTTTTGT------CTTGTGTTTC 1276

Qy   1201 TCTAGTTTTGTGAGGTTGACTTAATTATATAAAATATTGTTTTTGTTTGTTTCATGT 1260
            ||||||||||||||||||||||| |||||||||||||      |||||||||||||||||
Db   1277 TTTAGTTTTGTGAAGTTTAATTGACTTATTTATA----TAAATTTTTTTTGTTTCATAT 1332

Qy   1261 GAATGAGGCTCTAGGCAGGACCTGTGGCCAAGTTCTAGTAGCTGTTTATCGTGTGTAG 1320
            ||||||||| ||||||| ||||||||||||||||||| ||||| ||| | |||||||||
Db   1333 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTGTATGTGTCCGTGGTAG 1392

Qy   1321 GACTGTAGAACTGTGAGAGGAGAAACTGAACATTCCAGAATGTGTGAGTAGGTAAATTGAATAAA 1380
            |||||||||       ||||||||||||||||||||||
Db   1393 GACTGTAGAA------AAGGGAACTGAACATTCCAGAGCCGTGTTTATGCATA 1444

Db   1381 GCTAGCCGTGATCCTCAGCTGTGCTGCATA 1411
            ||||||||||||||||||||||||||||||
Db   1445 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1475
```

```
RESULT 13
PCT-US94-06380-2
; Sequence 2, Application PC/TUS9406380
GENERAL INFORMATION:
   APPLICANT: Chuntharapai, Anan
   APPLICANT: Lee, James
   APPLICANT: Hebert, Caroline
   APPLICANT: K. Jin Kim
   APPLICANT: Genentech, Inc.
   TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATORY DISRODERS
   NUMBER OF SEQUENCES: 3
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Genentech, Inc.
      STREET: 460 Point San Bruno Blvd
      CITY: South San Francisco
      STATE: California
      COUNTRY: USA
      ZIP: 94080
   COMPUTER READABLE FORM:
      MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: patin (Genentech)
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: PCT/US94/06380
      FILING DATE:
      CLASSIFICATION:
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 08/076093
      FILING DATE: 11-JUN-1993
   ATTORNEY/AGENT INFORMATION:
      NAME: Love, Richard B.
      REGISTRATION NUMBER: 34,659
      REFERRENCE/DOCKET NUMBER: 706P2P1
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: 415/225-5530
      TELEFAX: 415/952-9881
      TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 1737 bases
      TYPE: nucleic acid
```

```
;     STRANDEDNESS: single
;     TOPOLOGY: linear
PCT-US94-06380-2

Query Match       55.1%;    Score 933.4;   DB 5;   Length 1737;
Best Local Similarity   81.8%;   Pred. No. 1e-210;
Matches 1154;   Conservative   0;   Mismatches 216;   Indels 41;   Gaps    5;

Qy      1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG 60
            |||||||||||||||| ||||||||| ||||| |||||||||||||||||||||||||
Db    106 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 165

Qy     61 GAACCCTGCTTCCGGGATGAGAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
            |||||||||||| | |||||||| ||||||  |||||||||||||||||||||||
Db    166 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC 225

Qy    121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGTGTGATCCTGGTCATGGGTTAC 180
            |||||||||||| ||||||||| ||||||||||||||||||||||||||
Db    226 TCCATCATCTTCTTAACTGGCATATGTGGGCAATGGTGTCATCCTGGTCATGGGTTAC 285

Qy    181 CAGAGAAGCTAAGGAGCATGACGGACAAGTACCGGCGTCACCTGTCAGTGGCTGACCTC 240
            |||||||||| ||||||||||||||| ||||||||| ||||||||||||||||||||
Db    286 CAGAGAAACTGAGAAGCATGACGGACAAGTACAGGCGTCACCTGTCACCTGTCAAACTGTCAGTGGCCGACCTC 345

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCCATTCGACTGGTACTTTGGG 300
            ||||||||||||||||| ||||||||||||| |||||||||||||||||||
Db    346 CTCTTTGTCATCACGCTCCCTTCCCTTCGGGCAGTTGATGCCATGGCCATTCGACTGGTACTTTGGG 405

Qy    301 AAATTTTGTGTAAGGCTGTCCATATCATCATCTACACTGTCAAGCTTCACAGCAGCGTTCTC 360
            ||||||||| |||||||||| |||||| |||||||||||||||||||||||||||
Db    406 AACTTCCTATGCAAGGCAGTCCATGTCATCTCATCTCACACAGTCATCCTACAGCAGTGTCCTC 465

Qy    361 ATCCTGGCCTTCATCAGCCTGGACCGGTACTCCGCCTTCCGGCAGTCCAGCCCCAGTCAA 420
            |||||||||||||||||||||||||| |||||||||||||||||||||||
Db    466 ATCCTGGCCTTCATCAGCTTCATCGGTGGACCGGTACCTGGCCATCCGCACCCAGCCCCTC 525

Qy    421 AGGCCAAGGAAAACTGCTGGCTGAAAAAGGCAGTCTATGTGGCGTCTGGAGATCCAGCCCTC 480
            ||| ||||||| ||||||||||||||| ||||||||| |||||||||||||||||
Db    526 AGGCCAAGGAAGCTGTTGGCGTGAAAGTGTCTCATGTGGCGTCTGGAGATCCCTGCCCCTC 585

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGGG 540
            |||||||||| ||||||||||||||| ||||||||||
Db    586 CTGCTGACATTCCCGACTTCATCTTTGCCAACG-----------TCAGTGAGGCA 630

Qy    541 GATGACAGGTACATCTGTGACGCCGCCTTTACCCCGGATAGCCTGTGGATGGTGGTGTTTCAA 600
            |||||||||||||||||||||||||||||||||| ||||||| ||||||||| |||| ||
Db    631 GATGACAGATATATCTGTGACCGCGTTCTCTACCCCCAATGACTGTTGTGGGTGGTGTTGTTCCAG 690

Qy    601 TTCCAGCATATAAATGGTGGGTCTCATCCTTCCATCCTCTTCCTGTTACTGC 660
            ||||| |||| ||||||||||| |||||||| |||||| |||||||| ||
Db    691 TTTCAGCCACATCATGGTTGGCCTTATCCTCTGTATTGTCATCCTGTCCTGTTATTGC 750

Qy    661 ATCATCATCCTAAGCTGTCACACTCCAAGGGCGCCACCAGAGCGGCAAGGACG 720
            ||||||||||| ||||||||||||||||||||| ||||||||||| |||||
Db    751 ATTATCATCTCCAAGTGTCACACTCCAAGGGCGCCACCAGAGCGCCCAAGACC 810

Qy    721 ACAGTGCATTCATCCTCAGCTTTCTTTGCCTGCTGGTGGCCATATTATGTGGGGATCAGC 780
            |||||||||||||||||||||||||||||||||||||||||||
Db    811 ACAGTGCATCATCCTCATCCTCGGCTTTCGCCTGTTGGCTGCTTACTACATTGGGATCAGC 870

Qy    781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCCGAGAGCATTGTG 840
            |||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db    871 ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGATGTTTGAGAACACTGTG 930

Qy    841 CACAAGTGGAATCCATCACAGCAGAGGCCCTGCCTTCTTGCCTGAACCCCATC 900
            ||||||||| |||||||||||||||||||||| |||||||| ||||||||
Db    931 CACAAGTGGATTCCATCACCACCGGAGCCCTAGCTTTCTTCCTCACTGTTGTCGAACCCCATC 990

Qy    901 CTCTATGCCTTCCTCGGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG 960
            |||||| |||||||| ||||||||||||||||| ||||||||||||||||||
Db    991 CTCTATGCTTTCCTTGGAGCCCAAATTTAAAAACCCTCTGCCCCAGCACGCACCACCTCTGTG 1050

Qy    961 AGCAGAGAGGCTCCAGCCTCCAAGATCCTTCAAGATCCTTCCAAGGAAAGGAAAGCGGG...GGACTCCTTCCGCGTC 1020
```

```
Db   1051 AGCAGAGGGGTCCAGCCTCCAGATAATCCTCCTCCAGTTTCACTCCAGCTAAGGAAAGCGAGGTGGACATTCATCTGTT 1110
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAAGGAAAGCGAGGTGGACATTCATCTGTT

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAAGGAAAGGCGAGGTGGACATTCATCTGTT 1080
            ||||||| |||||||||||||| |||||||||||
Db   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAA-------CACAGATGTAAAA 1080

Db   1111 TCCACTGAGTCTGAGTCTTCAAGTTTTCACTCCAGCTAA-------CACAGATGTAAAA 1162

Qy   1081 AATATATATATTATGATAAAGAACTTTTTATGTTACACATTTTCCAGATATATAAGAG 1140
            ||||||||||||||||||||| |||||||||||||||||||||||||
Db   1163 GACTTTTTTATACGATAAATAACTTTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAAG 1222

Qy   1141 ACTGACCAGTCTGTACAGTGTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
            |||||||||||||||||||||||||||||||||||||
Db   1223 ACTGACCCAATATTGTACAGTTTTTTTATTGCTTGTTGTTGATTTTTGT-------CTTGTGTTTC 1276

Qy   1201 TCTAGTTTTGTGAGGTTGACTTAACTTAATTATATAATAAATATTGTTTTTTGTTTCATGT 1260
            ||||||||||||||||||||| |||| |||||||| |||||||
Db   1277 TTTAGTTTTGTGAAGTTAATTGACTTATTTATA-----TAAATTTTTTTGTTTCATAT 1332

Qy   1261 GAATGAGCGTCTAGGCAGGACCTGTGGCCAAGTCCTTAGTAGCTGTGTTTATCTGTGTGTAG 1320
            |||||||||||||||||||||||||||||||||||||||| ||||||
Db   1333 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTGTGTTAGTGTCCGTGGTAG 1392

Qy   1321 GACTGTAGAACTGTAGAGGAGGAGAAACTGAACATTCCAGAAATGTGTGTAAATTGAATAAA 1380
            ||||||||| |||||||||||| |||||||||||||||||||
Db   1393 GACTGTGAGAA-------AAGGGAACTGAACATTCCAGAGCGTGTTTAGTGAATCACGTAAA 1444

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
            |||||||||||||||||||||||||||||
Db   1445 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1475
```

```
RESULT 14
US-09-016-434-1235
; Sequence 1235, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
;    APPLICANT: Janice Au-Young
;    APPLICANT: Jeffrey J. Seilhamer
;    TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES: 1490
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;       STREET: 3174 PORTER DRIVE
;       CITY: PALO ALTO
;       STATE: CALIFORNIA
;       COUNTRY: USA
;       ZIP: 94304
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/016,434
;       FILING DATE: HEREWITH
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;    CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Zeller, Karen J.
;       REGISTRATION NUMBER: 37,071
;       REFERENCE/DOCKET NUMBER: PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (650) 855-0555
;       TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1235:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: 1225 base pairs
;       TYPE: nucleic acid
```

Query Match   50.5%; Score 855.6; DB 4; Length 1225;
Best Local Similarity 83.3%; Pred. No. 2e-192;
Matches 1005; Conservative   0; Mismatches 179; Indels 22; Gaps   2;

```
Qy    1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG   60
          ||||||||||| ||||||||  | || ||| |||| ||  ||||||||||||||||| |||
Db   41 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG  100

Qy   61 GAACCCTGCTTCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC   120
          |||||||| || | | |||||| |  ||||||||||  ||||||||||||||||||||
Db  101 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTAC   160

Qy  121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATGGTGATCCTGGTCATGGGTTAC   180
          | |||||||||||| ||||||||| || ||||||| ||| ||||||||||||||||||
Db  161 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGATTGGTCATCCTGGTCATGGGTTAC   220

Qy  181 CAGAAGAAGTAAGGACGATGACGGACAAGTACCCGGCTGCACCTGCACCTGTCAGTGCTGACCTC  240
          |||||||| | ||| ||||||||||| ||||| |||||| |||| ||||| |||||||||||
Db  221 CAGAAGAAACTGAGAAGCATGACGGACACAGTACCACGGCTCACCTGTCAGTGGCCGACCTC   280

Qy  241 CTCTTTGTCATCACACTCCCCTTCTCGGCAGTTGATGCCATGGCCATCTGACTGTACTTTGGG  300
          |||||||||||||| ||||| ||||| |||||||||||| ||||| ||| |||||||||||||
Db  281 CTCTTTGTCATCACGTCCCTTCTCTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG     340

Qy  301 AAATTTTTGTGTAAGCTGTCCATATCATCTACACGTCAACCTCTACAGCAGCGTTCTC    360
          || ||| | ||| ||||||||||| |||||||||| |   || ||||||| ||| ||
Db  341 AACTTCCTATGCAAGGCAGTCCATGTCATCTACACCAGTCTTACAGCAGTGTCCTC       400

Qy  361 ATCCTGGCCTTCATCAGCCTGACCGGTACCTGCCATTGTCCAGCCACCAACAGTCAA    420
          ||||||||||||||||| ||| |||||||||| || |||||||||||||||||||| |
Db  401 ATCCTGGCCTTCATCAGTCTGGACCGGTACCTGGCACCTGTCCAGCCACCAACAGTCAG     460

Qy  421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCAGCCCTC  480
          |||||||||||||| |||||| ||||||  ||||||||  ||| ||||||||| |||||||
Db  461 AGGCCAAGGAAGCTGTTGGCTGGTGAAAAGGTGTCTATGTTGCTGTCGGATCCCTGCCCTC   520

Qy  481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACACATCAGTCAGGGG  540
          || |||||||| ||| |||||||||||||  | |       |||| |  |||||
Db  521 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG--------TCAGTGAGGCA            565

Qy  541 GATGACAGGTACATCTGACCGCCTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA   600
          ||||||| | ||| |||||||||||  ||||  | ||||| |||||||||| ||||||
Db  566 GATGACAGAGATATATCTGACCGCCTTCTACCCCACCCAATGACTTGTGGGTGGTGTTTCCAG   625

Qy  601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTGTTACTGC   660
          || |||||| | |||| ||| ||||| |||  | | |||||||||||| |||| |||||
Db  626 TTTCAGCACACATCATGGCCTTATCTGCGTATTGTCATCCTGCTCGTGCTCATTGC       685

Qy  661 ATCATCATCTTCAAGCTGTCACACTCCAAGGCCACCAGAGGCGGCAAGGCCCTCAAGACG  720
          || |||||| ||||||||||||||||||||| |||||||||||| |||||||||||||||
Db  686 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAGGCGCAAGGCCCTCAAGACC   745

Qy  721 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGGCTGCCATATTATGTGGGGATCAGC  780
          |||||||||||||||||| ||||||||  |  |||||||||| || ||| | |||||||||
Db  746 ACAGTCATCCTCATCCTGGCTTTCTTCTCGCCTTGGCTGCCTTACTACACATTGGGATCAGC   805

Qy  781 ATCGACTTCTTCATCCTTTGGGAGTCATCAAGCAAGGATGGACTTCGAGAGCATTGTG   840
          ||||||||||||||||| ||||| | |||||||||||  |  | | |||| |||||||
Db  806 ATCGACTTCTTCATCCTTCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG    865

Qy  841 CACAAGTGGATCCTCATCACAGAGAGGCCCTTCTTGTGCTTGGCCTGAACCCCATC    900
          ||||||||||| ||||||||| ||||| |   |  ||| ||  ||||||||||||||
Db  866 CACAAGTGGATTTCATCACCGAGGCCCTAGCTTTCTTCCACTGTTGTCTGAACCCCATC    925

Qy  901 CTCTATGCCTTCCTCGGGGGCAAGTCTGCCCAGCATGCCACTCCACTCCAATGCCATG   960
          |||||||| |||||||||||||||| |||||||||||||||||||||||  |
```

```
Db  926 CTCTATGCTTCCTTGGAGCCAAATTTAAAACCTCTGCCCAGCAGCGCACCTCACCTCTGTG  985
          ||||||||  ||||| |||||||||  ||||||||||||||||  | |||  || ||||||
Qy  961 AGCAGAGCTCCAGCCTCAAGATCCTTTCCAAAGGAAGCGGGGTGGACACTCTTCCGTC    1020
          |||||||  ||||||||| |||||| |||||||||||| | |||||||| |  |  | ||
Db  986 AGCAGAGGGTCCAGCCTCCAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCGTT   1045

Qy 1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCTTATGCAAAGACTTATAT  1080
          ||||| |||||  |  || |||||| |||||||||||||   | | |||  |
Db 1046 TCCACTGAGTCTGAGTCTTCAGTTTTCACTCCAGCTAACACAGATGTAAA------AG  1098

Qy 1081 AATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG    1140
          | |||||| | | | ||||| |  |||||| | |||| |||||| |||||||| |||
Db 1099 ACTTTTTTTATACGATAAATAACTTTTTTTTAAGTTACACATTTTCAGATATAAAAG   1158

Qy 1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC  1200
          |||||||| | | |||||||||||  |   | |  | |||| |||  ||| |||| |
Db 1159 ACTGACCAATATTGTACAGTTTTTTTATTGCTTGTTGTTTTTGCTTGTTTTCTTTAGT    1218

Qy 1201 TCTAGT 1206
          | | ||
Db 1219 TTTTGT 1224
```

RESULT 15
US-08-153-848-45

```
Query Match           49.2%;   Score 834;  DB 1;  Length 1317;
Best Local Similarity 84.1%;   Pred. No. 2.5e-187;
Matches 974;  Conservative  0;  Mismatches 160;  Indels  24;  Gaps  2;

Qy    1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG   60
Db  171 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTACGGGGACTATGACTCCATGAAG  230

Qy   61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC   120
Db  231 GAACCCTGTTTCCGTGAAGAGAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTAC  290

Qy  121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATGGTGATCCTGGTCATGGGTTAC   180
Db  291 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC  350

Qy  181 CAGAAGAAGCTAAGGAGCATGACGACGAACAAGTACCGGCTGCACCTGCACCTC       240
Db  351 CAGAGAAAACTGAGAGCATGACGGACAAGTACGGCTGCACCTGTCAGTGGCCGACCTC    410

Qy  241 CTCTTTGTCATCACACTCCCCTTCTGGCAGTGATGCCATGGCCATGGACTTTGGG      300
Db  411 CTCTTTGTCATCACGCTTCCCTTCTGTGGGCAGTTGATGCCGTGGCAAACTGGTACTTTGGG 470

Qy  301 AAATTTTTGTGTAAGGCTGTCCATATCATCTTACACTGTCAACCTCTACAGCAGCGTTCTC 360
Db  471 AACTTCCTATGCAAGGCAGTCCATGTCATCTTCACACAGTCAACCTCTACAGCAGTGTCCTC 530

Qy  361 ATCCTGGCCTTCATCAGCGCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAA 420
Db  531 ATCCTGGCCTTCATCAGTCTGGACCGGTACCTGGCCATCGTCCACGCCACCAACAGTCAG   590

Qy  421 AGGCCAAGGAAACTGCTGGCTGAAAAGCAGTCATGTGGGCGTCTGGATCCCAGCCCTC    480
Db  591 AGGCCCAAGGAAGCTGTTGGCTGAAAAGGTGTCTATGTTGGCGTCTGGATCCCCTGCCCCTC 650

Qy  481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG   540
Db  651 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG----------TCAGTGAGGCA       695

Qy  541 GATGACAGTACATCTGTGACCGCCCTTACCCGATAGCCTGTGGATGGTGGTGTTCAA     600
Db  696 GATGACAGAGTATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGGTGGTGTTCCAG  755

Qy  601 TTTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTTCCTGTTACTGC  660
Db  756 TTTCAGCACACATCATGGTTGGCCTTATCCTGCCTGGTATTGTCATCCTGTCCTGCTATTGC 815

Qy  661 ATCATCATCCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGGCAAGGCCCTCAAGACG 720
Db  816 ATTATCATCTCCCAAGCTGTCACACTCCAAGGGCCACCAGAAGCGGCAAGGCCCTCAAGACC 875

Qy  721 ACAGTCATCTCCATCAGCTTTCTTTTGCCTGCCTGCCATATTATGTGGGGATCAGC     780
Db  876 ACAGTCATCCATCATCCTCTCAGCTTTCTTTCGCCTGTTGGCTGCCTTACTACACATTGGGATCAGC 935

Qy  781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCAGAGAGCATTGTG  840
Db  936 ATCGACTCCTTCATCCTTCCTGGAAATCATCAAGCAGGGTGTGAGTTTGAGAACACTGTG   995

Qy  841 CACAAGTGGATCTCCATCACAGAGAGGCCCTGCCTTCTTCTTCCTGCCTGACCCCATC    900
Db  996 CACAAGTGGATTTCCATCACCCGAGGCCCTAGCTTCTTCTTCCACTGTTGCTGAACCCCATC 1055

Qy  901 CTCTATGCCTTCCTGGGGGCCAAGTCAAAGGCTCAAAAGCTCGCACTCAACCTCCATG    960
Db 1056 CTCTATGCTTCTCCTTGGAGCCAAATTAAAACCTCGCCCAGCACGCACCTTCGTG      1115

Qy  961 AGCAGAGGCTCCAGCCTCAAGATCCTTTCAAAGGAAGGAAAGGGGGTGGACTCTTCCGTC 1020
Db 1116 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTT  1175

Qy 1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT  1080
Db 1176 TCCACTGAGTCTGAGTCTTCAGTTTTCACTCCAGCTAACACAGATGTAAA-----------1226

Qy 1081 AATATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAG   1140
Db 1227 AGACTTTTTTATACGATAAATAACCTTTTTTTAAGTTACACATTTTCAGATATAAAG    1286

Qy 1141 ACTGACCAGTCTTGTACA 1158
Db 1287 ACTGACCAATATTGAAAA 1304

Search completed: May 17, 2004, 16:15:55
Job time : 145.63 secs
```

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      May 17, 2004, 12:11:00 ; Search time 829.19 Seconds
                                      (without alignments)
                              9271.041 Million cell updates/sec

Title:           US-09-367-052-7
Perfect score:   1694
Sequence:        1 atatacacttctgataacta.......taaaagtcatgatataactt 1694

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2947324 seqs, 226902451 5 residues

Total number of hits satisfying chosen parameters:      5894648

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
          1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
          2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
          3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
          4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
          5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
          6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
          7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
          8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
          9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
         10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
         11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
         12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
         13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
         14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
         15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
         16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
         17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
         18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
         19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	933.8	55.1	5161	16	US-10-160-401-1	Sequence 1, Appli
2	933.4	55.1	1664	17	US-10-641-643-1213	Sequence 1213, Ap
3	933.4	55.1	1679	13	US-10-211-462-80	Sequence 9, Appli
4	933.4	55.1	1679	13	US-10-181-906-9	Sequence 9, Appli
5	933.4	55.1	1679	13	US-10-342-887-912	Sequence 912, App
6	933.4	55.1	1679	13	US-10-151-274-14	Sequence 14, Appl
7	933.4	55.1	1679	13	US-10-172-118-912	Sequence 912, App
8	933.4	55.1	1679	13	US-10-170-385-332	Sequence 332, App
9	933.4	55.1	1679	15	US-10-225-660-58	Sequence 75, Appl
10	933.4	55.1	1679	15	US-10-021-660-58	Sequence 58, Appl
11	933.4	55.1	1679	16	US-10-341-434-177	Sequence 177, App
12	933.4	55.1	1711	10	US-09-971-392-20	Sequence 20, Appl
13	933.4	55.1	1711	15	US-10-101-510-459	Sequence 459, App
14	933.4	55.1	1737	9	US-09-104-063-3	Sequence 3, Appli

15	933.4	55.1	1737	13	US-10-666-689-3	Sequence 3, Appli
16	931.4	55.0	1670	9	US-09-880-107-2143	Sequence 2143, Ap
17	931.4	55.0	1670	10	US-09-960-706-636	Sequence 636, App
18	931.4	55.0	1670	16	US-10-372-683-1	Sequence 1, Appli
19	931.4	55.0	1670	16	US-10-440-464-77	Sequence 77, Appl
20	855.6	50.5	1225	15	US-10-101-510-741	Sequence 741, App
21	855.6	50.5	1225	16	US-10-305-720-1235	Sequence 1235, Ap
22	843.6	49.8	1224	15	US-10-101-510-674	Sequence 674, App
23	806.2	47.6	1059	13	US-09-813-651B-84	Sequence 84, Appl
24	806.2	47.6	1059	15	US-10-014-322A-125	Sequence 125, App
25	806.2	47.6	1059	16	US-10-160-401-2	Sequence 2, Appli
26	804.6	47.5	1102	9	US-09-870-759-143	Sequence 143, App
27	804.6	47.5	1102	10	US-09-751-708A-143	Sequence 143, App
28	801.4	47.3	1059	11	US-09-826-509-484	Sequence 484, App
29	553.6	32.7	1902	9	US-09-953-692-1	Sequence 1, Appli
30	553.6	32.7	1902	9	US-09-953-717-1	Sequence 1, Appli
31	530	31.3	1076	15	US-10-017-161-789	Sequence 789, App
32	530	31.3	1076	16	US-10-292-798-687	Sequence 687, App
33	345.2	20.4	507	13	US-10-085-783A-14955	Sequence 14955, A
34	345.2	20.4	507	16	US-10-242-535A-14955	Sequence 14955, A
35	303.2	17.9	421	9	US-09-795-692-2611	Sequence 2611, Ap
36	303.2	17.9	421	15	US-10-040-862-2611	Sequence 2611, Ap
37	303.2	17.9	421	16	US-10-057-475B-2611	Sequence 2611, Ap
38	303.2	17.9	421	16	US-10-154-884B-2611	Sequence 2611, Ap
39	287	16.9	487	9	US-09-795-692-8936	Sequence 8936, Ap
40	287	16.9	487	15	US-10-040-862-8936	Sequence 8936, Ap
41	287	16.9	487	16	US-10-057-475B-8936	Sequence 8936, Ap
42	287	16.9	487	16	US-10-154-884B-8936	Sequence 8936, Ap
43	198.4	11.7	1119	12	US-09-997-722-162	Sequence 162, App
44	198.4	11.7	1119	15	US-10-251-385-65	Sequence 65, Appl
45	198.4	11.7	1679	9	US-09-104-063-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-10-160-401-1
; Sequence 1, Application US/10160401
; Publication No. US20030207281A1
; GENERAL INFORMATION:
; APPLICANT: Genaissance Pharmaceuticals, Inc.
; APPLICANT: Bentivegna, Steven C.
; APPLICANT: Bieglecki, Karyn M.
; APPLICANT: Koshy, Beena
; APPLICANT: Monroe, Glen
; APPLICANT: Rounds, Eileen
; TITLE OF INVENTION: HAPLOTYPES OF THE CXCR4 GENE
; FILE REFERENCE: MWH-0121US
; CURRENT APPLICATION NUMBER: US/10/160,401
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: PCT/US01/12268
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,025
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 1
;   LENGTH: 5161
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (3118)..(3118)
;   OTHER INFORMATION: PS1: polymorphic base thymine or adenine
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (3583)..(3583)
;   OTHER INFORMATION: PS2: polymorphic base cytosine or thymine
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (3952)..(3952)
;   OTHER INFORMATION: PS3: polymorphic base cytosine or thymine
```

```
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (4030)..(4030)
;   OTHER INFORMATION: PS4: polymorphic base cytosine or thymine
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (4262)..(4262)
;   OTHER INFORMATION: PS5: polymorphic base adenine or thymine
US-10-160-401-1

Query Match     55.1%;  Score 933.8;  DB 16;  Length 5161;
Best Local Similarity   81.6%;  Pred. No. 3.7e-236;
Matches 1152;  Conservative   3;  Mismatches 215;  Indels 41;  Gaps   5;

Qy      1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGTCTGAGACTATGACTCCAACAAG 60
            |||||||||||||||| ||||| ||||||||||||||| |||||||||||||||||
Db   3185 ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 3244

Qy     61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCACCANCTAC 120
            |||||||||||| |||| ||||||||||||| |||| ||||| |||||||||||| |||
Db   3245 GAACCCTGTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCCACCATCTAC 3304

Qy    121 TTCATCATCTTCTTGACTCGGCATAGTCCGCCAATGGATCGATCCTGTCTCATGGGTTAC 180
            || |||||||||||||||| ||||| ||||||| |||||||| |||||||||||||||||
Db   3305 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGTCATGGGTTAC 3364

Qy    181 CAGAGAGAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
            ||||||||| ||||| |||||||||||||||| ||||| ||||||| |||||| |||||
Db   3365 CAGAGAGAACTGAGAAGCATGACAGCAGTACACAGGCTGCACCTGTCAGTGGCCGACCTC 3424

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTCTTTTGGG 300
            || |||||||||||||||||||||||||| ||||||||||| ||||| |||||||||||
Db   3425 CTCTTTGTCATCACAGTCCCCTTCTGGGCAGTTGATGCCATGGCTGACTCTTTTGGG 3484

Qy    301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCCGTTCTC 360
            |||||||||| ||||||||||||| ||||||||||||||||||||||||||| |||||||
Db   3485 AACTTCCATGCAAGGCAGTCCATCCATGTCATCTACACACAGTCAACCTCTACAGCAGTGTCCTC 3544

Qy    361 ATCCTGGCCTTCATCAGCCTGACGCCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAA 420
            |||||||||||||||||||||| |||||||| |||||||||| |||||||||||||||||
Db   3545 ATCCTGGCCTTCATCAGCTCGGACCGCTACCTGGCCATYGTCCACGCCACCAACAGTCAG 3604

Qy    421 AGGCCAAGGAAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGATCCCAGCCCTC 480
            ||||||||||||| |||||||||||||||||||||| |||||||||||||||||||||||
Db   3605 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTGGGTCTGATCCCTGCCCTC 3664

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCAGGGGACATCAGTCAGGGG 540
            || |||||||| || |||||| ||||||||||||||||| |||| || |
Db   3665 CTGGTCGACTATTCCCGACTTCATCTTTCATCCTTTGCCAACG-----------TCAGTGAGGCA 3709

Qy    541 GATGACAGTACATCGTGACCACTCTGACCGCCCTTACCCGTGGATGGTGGTGGTGTTTCAA 600
            ||||||| |||||||| |||||||| || |||||||||| ||||| |||||||| |||||
Db   3710 GATGACAGATATATCGTGACCGCTTCTACCCCAATGACTTGTGGGGTGTTGTTCCAG 3769

Qy    601 TTCCAGCATATAATGGTGGGTCATCATCCTGCCCGGCCATCGTCATCCTCCTGTTACTGC 660
            ||||| | |||| |||||||| || ||||||||| |||||||||||||||||| ||||||
Db   3770 TTTCAGCACCATCATGGTTGGCCTTATCACCTGGCCTATTGTCATCGTCCTGCTCATTGC 3829

Qy    661 ATCATCATCCTCTAAGGCTGTCACACTCCAAGGGCCCACCAAGGCCCTTCAAGACG 720
            ||||||||||| | |||||||||||| ||||||||||||||| |||||||||||||
Db   3830 ATTATCATCCTCCCCAGCTGTCACACTCCAAGAGCGCCACCAAGGCCCTCAAGACC 3889

Qy    721 ACAGTCATCCTCATCCTCAGCTTTCTTTGCCTGCCTGGCCCATATTATGTGGGGATCAGC 780
            |||||| |||||||||||||||||| |||| ||||||||||| || || |||||||||||
Db   3890 ACAGTCATCCTCATCCTTTGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCAGTTGTG 3949

Qy    781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGAGGTGTGAGTTTGAGAACACTGTG 840
            || |||||||||||||||| | |||| ||| ||| || |||||| || ||| || |||||
Db   3950 ATYGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 4009

Qy    841 CACAAGTGGATCTCCATCACAGAGACCCCCGCCTTCTTCCACTGTTGCCTGAACCCCCATC 900
            ||| |||||| ||| |||||||||||| |||| |||||||||||||||||||||||||||
Db   4010 CACAAGTGGATTTCCATCACACYGAGGCCCCTAGCTTTCTTCCACTGTTGTCTGAACCCCCATC 4069
```

```
Qy    901 CTCTATGCCTTCCTCGGGCCAAGTTCAAAAGCTCGCCAGCATGCACTCAACTCCATG 960
            |||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||
Db   4070 CTCTATGCTTCCTTGGAGCCAAATTAAAACCTCGCCCAGCACGCACCACCTCACCTCTGTG 4129

Qy    961 AGCAGAGGCTCCAGCCTCAGATCCTTCCAAAGGAAAGCGGGTGGACACTCTTCCGTC 1020
            ||||||||||||||||||||||||||| |||||||||| |||||||||||||||||||
Db   4130 AGCAGAGGGTCCAGCCTCAGATCCTCTCCAAAGGAAAGGAGGGTGGACATTCATCTGTT 4189

Qy   1021 TCCACCGAGTCAGAATCCTCCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
            ||||||||||||||||| |||||||||  ||||||||
Db   4190 TCCACTGAGTCTGAGTCTCAGTTTCACTCCCAGCTAA---------CACGATGTAAAA 4241

Qy   1081 AATATATATATATATGATAAGAACTTTTATGTTACACATTTCCAGATATAAGAG 1140
            || |   |  |  |  | | || |||||  |||||||||||||| ||||| ||||
Db   4242 GACTTTTTTTTATACGATAAWTAACCTTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 4301

Qy   1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTAATTGACTGTGTGGAGTTATGTTCC 1200
            ||||| ||||| |||||||||  | |||||| ||||| ||| |          ||||||
Db   4302 ACTGACCAAATTGTACAGTTTTTATTGCTTGTTGGATTTTTGT-------CTTGTGTTTC 4355

Qy   1201 TCTAGTTTTTGTGAGGTTGACTTAATTATATAATAAATATTGTTTTTGTTCATGT 1260
             | ||||||||||| ||| ||| || ||||||||||
Db   4356 TTTAGTTTTGTGAAGTTAATTGACTTATTATA----TAAATTTTTTTGTTTCATAT 4411

Qy   1261 GAATGAGCGTCAGGCAGGACCGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
            ||||||| |||||||||||||| |||||| |||||||| || |||||||| ||| |||||
Db   4412 TGATGTGTGTCTAGGCAGGACCTGTGGCCCAAGTTCTTAGTTGCTGTTATGTCTCGTGGTAG 4471

Qy   1321 GACTGTGAGAACTGTAGAGAGAAGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
            |||||||||| ||||||||      |||||||||||||||||| |||||| ||| |||
Db   4472 GACTGTAGAA--------AAGGGAACTGAACATTCCAGAGTGTAGTAGTGAATCACGTAAA 4523

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
            ||||| | ||||||| ||||||  |||||||
Db   4524 GCTAGAAATGATCCCCAGCTGTTTATGCATA 4554


RESULT 2
US-10-641-643-1213
;   Sequence 1213, Application US/10641643
;   Publication No. US20040077003A1
;   GENERAL INFORMATION:
;       APPLICANT: Cocks, Benjamin G.
;                  Susan G. Stuart
;                  Jeffrey J. Seilhamer
;       TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
;                           GENE EXPRESSION
;       NUMBER OF SEQUENCES: 1508
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;           STREET: 3174 PORTER DRIVE
;           CITY: PALO ALTO
;           STATE: CALIFORNIA
;           COUNTRY: USA
;           ZIP: 94304
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/10/641,643
;           FILING DATE: 14-Aug-2003
;           CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: <Unknown>
;           FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Zeller, Karen J.
;           REGISTRATION NUMBER: 37,071
;           REFERENCE/DOCKET NUMBER: PA-0001 US
;       TELECOMMUNICATION INFORMATION:
```

```
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
;   INFORMATION FOR SEQ ID NO: 1213:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1664 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     LIBRARY: GENBANK
;     CLONE: g219868
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1213 :
US-10-641-643-1213

Query Match    55.1%;    Score 933.4;  DB 17;  Length 1664;
Best Local Similarity  81.8%;  Pred. No. 2.1e-236;
Matches 1154;  Conservative  0;  Mismatches 216;  Indels 41;  Gaps  5;
```

```
Qy      1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGAGAGACTATGACTCCAACAAG 60
            ||||||||| |||||||||| |||||  |||  ||||||  ||||| || ||||||||||
Db     92 ATATACACTTCAGATAACTACTACACCGAGGAAATGGGCTTCAGGGGACTTATGACTCCATGAAG 151

Qy     61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 120
            |||||||||||| ||||  ||||| |||||||| ||| |||| ||||||||||||||||
Db    152 GAACCCTGTTTCCGTGAAGAAAATGTAATTTCAATTCAATAAATCTTCCTGCCCACCATCTAC 211

Qy    121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGCAATGGATGATTGGTGATCCTGGTCATGGGTTAC 180
            |||||||||||  ||||||||||||||||||| |||| ||| |||| ||||||| ||||||||||||||
Db    212 TCCATCATCTTCTTCTAACTGGCATTGTGGGCAATGGCAATGGATGGTGATCCTCGTCATGGGTTAC 271

Qy    181 CAGAAGAAGCTAAGGAGCCAAGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
            |||||||||| |||||||||||||||| |||||| |||||||| |||||||||||| 
Db    272 CAGAAGAAACTGAGAAGCCATGACGAAGCATGACGAAGCAAGTACAGGGGCCGGCTGGTGCCTGACCTC 331

Qy    241 CTCTTTGTCATCACACTCCCCTTCGGGCAGTTGATGCCATGGCTGACTGACTTTGGG 300
            ||||||||||||||||||||||||||||||| |||||||| ||||  ||||||||||
Db    332 CTCTTTGTCATCACGCTTCCCTTCGGGCAGTTGATGCCGTGCCAAACTGTACTTTGGG 391

Qy    301 AAATTTTTGTGTAAGGCTGTGTCCATATCATCTACACACTGTCAACCTCTACAGCAGCGGTTCTC 360
              |||||||| |||||||| |||||| | ||||||||| |||||||||||||||||| |||||
Db    392 AACTTCCATGCAAGGCAGTCCATGTCCATGTCATCTACACAGTCAACCTCTACAGCAGTGTCCTC 451

Qy    361 ATCCTGGCCTTCATCAGCCTGGACCGGACCGGTACCTGCCATTGTGCCACGCCACCACAGTCAA 420
            ||||||||||||||||||| ||||  |||  ||||||||||||||||||||||| ||||||||
Db    452 ATCCTGGCCTTCATCAGTCTGGACCGGTTACCTGCCATCGTGCCACGCCACCAAACAGTCAG 511

Qy    421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTTCTATGTGGGCGTCTGGATCCCAGCCCTC 480
            |||||||||||||||||||  |||||||| ||||| |||||||| ||||||||||||||||
Db    512 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTGGGCTCTGGATCCCTGCCCTC 571

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGGGGACATCAGTCAGTCAGGGGG 540
            |||||||||||| ||||||||||||||||||||| |||||  ||||||     |||||||
Db    572 CTGCTGACTATTCCCGACCTTCATCTTTGCCAACG--------TCAGTGAGGCA 616

Qy    541 GATGACAGTCACATCTGTGACGCGGCCTTTACCCGATAGCCTGTGGATGGTGTGTTCAA 600
            ||||||| |||||||| |||  |||||||||||   |||| |||||| ||||||||||
Db    617 GATGACAGATATATCTGTGACCGCGTTCTACCCCAATGACTTGGGTGGTTGTGTTCCAG 676

Qy    601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCCATCCTCCTGTTACTGC 660
            ||||||||| |||  |||  |||  ||| |||||  ||||||||||||||||||
Db    677 TTTCAGCACATCATGGTTGGCCTTATCCTGCCCTGGTATTGTCATCCTGCTATTGC 736

Qy    661 ATCATCATCTCTAAGCTGTCACACTCCAAGAGGCCACCAGAAGGCCAGGCCCTCAAGACG 720
            || ||||| |||  | ||||||||||||||||||||||||||||||||||||||||||||
Db    737 ATTATCATCTCCAAGCTGTCACACTCCAAGAAGCGCAAGGGCCCTCAAGACC 796

Qy    721 ACAGTCATCCTCATCCTAGCTTTCTTTGCCGTCGCCGGCTCCGCATATTAATGTGGGGGATCAGC 780
            |||||||||||||||||| |||||||| ||||||||| |||||||||||||||||||||||||||
Db    797 ACAGTCATCCTCATCCTCTGCTTTCTTCTTCGCCGTTGGCCTTGCCCTTACTACACAATTGGGATCAGC 856

Qy    781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTTGACTTCGAGAGAGCATTGTG 840
            || |||||||||||||||||||||||||||||| |||||||||||||| |||||||||||||
Db    857 ATCGACTCCTTCATCCTTCCTGCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAGTTTGAGACACTGTG 916

Qy    841 CACAAGTGGATCTCCATCACAGAGAGCGCCCGCCTTCTTCCACTGTTGCCTGAACCCCATC 900
            |||||||||||| |||||  ||||||||||||||||||||||| ||||| |||||||||||
Db    917 CACAAGTGGATTCCATCACCGAGGCCCTAGCTTTCTTCCACTGTGTCTGAACCCCATC 976

Qy    901 CTCTATGCCTTCCTCGGGGCCAAGTCAAAAGTCTGGCCCAGCAGAGTCGACTCAACTCCATG 960
            |||||||||||||||| ||  ||  |||||  ||||||||||||| ||| |||||||||||
Db    977 CTCTATGCTTTCCTTGGAGCCCAAATTAAAACCTCGCCCAGCACACGCACACTCACCTCGTG 1036

Qy    961 AGCAGAGGCTCCAGCTCAAGATCCTTCCAAAGAAGCGGGTGACACTCTTCCGTC 1020
            |||||||||||||||||||| |||| |||||||||||||| ||||||||| |||
Db   1037 AGCAGAGGGTCCAGCCTCAAGATCCTCCCAAAGGAAGCGAGGTGGACATTCATCTGTT 1096

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGTTTCACTCCAGCTAA------CACAGATGTAAA 1080
            |||| |||||| ||||| ||||||||||||||||||| ||||||||||    ||||| ||||||
Db   1097 TCCACTGAGTCCTGAGTCTTCAAGTTTCACTCCAGTTTCACTCCAGCTAA CACAGATGTAAAG 1140

Qy   1081 AATATATATATATATATGATAAAGAACTTTTTATGTTACACATTTTCAGATATAAAG 1140
            |  || ||||||| || |||||||  |||| ||||  |||| |||||||| |||||||
Db   1141 GACTTTTTTATACGATAAATAACTTTTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1208

Qy   1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
            ||||||||||||||||||||||  |||||||          ||||||||||||||||
Db   1209 ACTGACCAATATTGTACAGTTTTTATTGCTTGTGTTGTGATTTTTGT------CTTGTGTTTC 1262

Qy   1201 TCTAGTTTTTGTGAGGTTGACTTAATTTATATAAATATTGTTTTTGTTGTTTCATGT 1260
            ||||||||| |||||||||| ||| ||||| ||||||||    |||||||||||||
Db   1263 TTTAGTTTTTGTGAAGTTTAATTGACTTTATTTATA----TAAATTTTTTTGTTTTTCATAT 1318

Qy   1261 GAATAGACGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
            |||||||| |||  ||| ||||| || ||||| ||  |||||| ||||| ||||| |||||
Db   1319 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTATGTCTTCGTGGTAG 1378

Qy   1321 GACTGTAGAACTGTAGAGAGGAGAAACTGAACATTCAGAATTGTGTGGTAAATTGAATAAA 1380
            ||||||||         |||||  ||||| ||  ||||||| || ||||||||||||
Db   1379 GACTGTAGAA------AAGGGAACTGAACATTCCAGAGCGTGTGAGTGTAGTTAATTAATCACGTAA 1411

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
            |||||| |||||||||||||| |||| |||||
Db   1431 GCTAGAAATGATCCCCAGCTGTTTTATGCATA 1461
```

```
RESULT 3
US-10-211-462-80
;  Sequence 80, Application US/10211462
;  Publication No. US2004003345A1
;  GENERAL INFORMATION:
;   APPLICANT: Murray, Richard
;   APPLICANT: Glynne, Richard
;   APPLICANT: Watson, Susan R.
;   APPLICANT: Aziz, Natasha
;   APPLICANT: Eos Biotechnology, Inc.
;   TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
;   TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
;   FILE REFERENCE: 018501-006200US
;   CURRENT APPLICATION NUMBER: US/10/211,462
;   CURRENT FILING DATE:  2003-02-13
;   PRIOR APPLICATION NUMBER: US 09/784,356
;   PRIOR FILING DATE: 2001-02-14
;   PRIOR APPLICATION NUMBER: US 09/791,390
;   PRIOR FILING DATE: 2001-02-22
;   PRIOR APPLICATION NUMBER: US 60/310,025
;   PRIOR FILING DATE: 2001-08-03
;   PRIOR APPLICATION NUMBER: US 60/334,244
;   PRIOR FILING DATE: 2001-11-29
;   NUMBER OF SEQ ID NOS: 230
;   SOFTWARE: PatentIn Ver. 2.1
;
;  SEQ ID NO 80
;  LENGTH: 1679
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-211-462-80
```

```
Query Match
Best Local Similarity   55.1%;   Score 933.4;   DB 13;   Length 1679;
Matches 1154;   Conservative   0;   Mismatches 216;   Indels   41;   Gaps   5;

Qy      1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG   60
          ||||||||||||| || ||| ||||||||||||||||||||||||| ||||||||| |
Db    104 ATATACACTTCAGATAACTACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG  163

Qy     61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC  120
          ||||||| ||||||||||||||||||| ||||| |||||| |||||||||||||||||
Db    164 GAACCCTGTTTCCGTGAGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC  223

Qy    121 TTCATCATCTTCTTGACTCGGCAATGAGTCGGCAATGGTGTGATCCTGGTCATGGGTTAC  180
          ||||||||||| ||||| ||| ||||| |||||| ||| ||||||||||| ||||||||||
Db    224 TCCATCATCTTCTTAACTGGGCAATGATTGGGCAATGGTGTGATCCTGGTCATGGGGTTAC  283

Qy    181 CAGAGAGAGCTAAGGACGTAAGGACGATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  240
          |||||||||||||||||||||||||||||||||||||| ||||| |||||||||| ||||||||||
Db    284 CAGAGAGAAACTGAGAAGCATGACGGACAAGTACAGGCGGGTGAACCTGTCAGTGGCCGACCTC  343

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCAGTGACTGGTACTTTGGG  300
          ||||||||||||||||||||||||||||||||||| |||||||| |||||||||||||
Db    344 CTCTTTGTCATCACGCTCCCCTTCTGGGCAGTTGATTCCGCAAACTGGTACTTTGGG  403

Qy    301 AAATTTTGTGTAAGGCTGTCTCATATATCACACTGTCAACCTCACAGCAGCGGTTCTC  360
          ||||| ||| |||| ||||||||| || || ||||||||||||||||| || || |||
Db    404 AACTTCCTATGCAAGGCAGTCCATGCATCTCATCACACAGTCAACCTCACAGCAGTGTCCTC  463

Qy    361 ATCCTGCCTTCATCAGCCTGACCCGGTACCTGCGCCATTGTCCAGCCACCAAACAGTCAA  420
          ||| ||| ||| |||||| ||||| ||||| || |||||||||||| |||||| |||||
Db    464 ATCCTGGCTTCATCAGTCTGGACCGCTACCTGCGGCATCGCGTCTGGATCCCAGCCCTC  523

Qy    421 AGGCCAAGGAAACTGCTGGCTGGAAAAGGCAGTGCTATGTGGGCGTCTGGATCCCAGCCCTC  480
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db    524 AGGCCAAGGAAACTGCTGGCTGGAAAAGGTGGTCTATGTGGGCGTCTGGATCCCTGCCCTC  583

Qy    481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCAGCGTCAGCATCAGTCAGGGG  540
          |||| |||||| ||||| || |||||||||||| ||||||||||  |||||||||||||
Db    584 CTGCTGACTATTCCCGACTTCATCTTTGCCAACGTCAGTGAGGCA  628

Qy    541 GATGACAGGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA  600
          |||||||||||| || ||||||||| ||| |||||||||||||||||||| |||||||||
Db    629 GATGACAGATATCTGTGACCGCCTTCTACCCCCAATGACTTGTGGGTGGTGTGTTCCAG  688

Qy    601 TTCCAGCATATAATGTGGGTGTCGTCATCATCCTGCCCGCGATCATCGTCATCCTCCTGTTACTGC  660
          |||| ||||| || |||||| |||||| ||| |||||||| ||||| |||||| |||||||||||||
Db    689 TTTCAGCAACATCATGGTTGGCCTTATCATCTTGCCTTATCCTGGTATTGTCATCCTGCTATTGC  748

Qy    661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAAGGGCCACCAAGGGCCAAGGCCCTCAAGACG  720
          ||| ||||||||| |||||| |||||||| ||| ||||||||||||||||||||||||| ||||||
Db    749 ATTATCATCTCCAAGCTGTCACACTCCAAGGGGCCAAGGGCCACCAAGGGCCAAGGCCCTCAAGACC  808

Qy    721 ACAGTCATCCTCATCCTAGCTTTCTTGCCTCGCTGCCCAATTATGTGGGATCAGC  780
          ||||||||||||||||||||||||| ||||||| ||||| ||||| ||||||||||
Db    809 ACAGTCATCCTCATCCTGGCTTTCTTCTGCCTCGTTGGCCTGCCTTACTACAATTGGGATCAGC  868

Qy    781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  840
          ||||||||||||||||| |||||||| || ||| ||||| |||||| |||||||  ||
Db    869 ATCGACTCCTTCATCCTTTGGGAAATCATCAAGCAAGGTGTGAGTTTGAGAACACTGTG  928

Qy    841 CACAAGTGGATCTCCATCACAGAGGCCCTGCCTTCTTCCCACTGTGCCTGAACCCCATC  900
          ||| |||||||| ||||||||||||||| |||||| ||||| |||||||||||||||||
Db    929 CACAAGTGGATTCCCATCACACCGAGCCCTGACTTCTTCCACTGTGTCTGCCAACCCCATC  988

Qy    901 CTCTATGCTTCCTCGGGGCCGAAGTCAAAAGCTCTGCCCAGCATGCATGCCAACTCCATG  960
          ||| ||||||||||| |||||||||||||| |||| ||||||||||| ||||| ||||||
Db    989 CTCTATGCTTTCCTTGGAGCCCAAGTTCAAAATCTGCCCAGCACCAACTCACCCCTGTG  1048

Qy    961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGAAAGCGGGGTGACACTCTTCCGTC  1020
          |||||||| ||||||||||||||||| |||||| || ||||||| || |||||||||
Db   1049 AGCAGAGGGTCCAGCCTCAAGATCCTCCCCAAAGGAAGCGGAGGTGAGCATTCATCTGTT  1108

Qy   1021 CTCTTTGTCATCACACTCCCCTTCACTCCAGCTAACCCTTATGCAAGAGACTTATAT  1080
          ||||| ||||||| ||||||||||||||||||| || ||||| |||||||||||||||
Db   1109 TCCACTGAGTCTGAGTCTTCAAGTTGATAAAGAAGAACTTTTTATGTTACACATTTTCAGATGTAAA  1160

Qy   1081 AATAATATATATAATATATGATAAAGAAACTTTTTTTTTTTTATGTTACACATTTTCAGATATAAGAG  1140
                                                        |||||||||||||| |||||||
Db   1161 GACTTTTTTTATACGATAAATAACTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG  1220

Qy   1141 ACTGACCAGTCTTGTACACAGTTTTTTTTTTTTTTTTTAATTGACTGTGGGAGTTTATGTCC  1200
                                                          |||||||||
Db   1221 ACTGACCAATATTGTACAGTTTTTTTATTGCTGTGTTTTTGT------CTTGTGTTTC  1274

Qy   1201 TCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAAATATTGTTTTTTGTTTGTTTCATGT  1260
          |||||| ||||||||||| |||||| ||  ||     || |||||| ||||||||| ||
Db   1275 TTTAGTTTTTGTGAAGTTTAATTGACTTATTTATA----TAAATTTTTTTGTTTTGTTTCATAT  1330

Qy   1261 GAATGAGCGTCAGGCAGGACCCTGGCCAAGTTCTTAGTAGCGTGTTTATCTGTGTGTAG  1320
          |||||||| |||| ||||||| |||| ||||||||||| |||| ||| ||||||||||
Db   1331 TGATGTGTGTCTAGGCAGGACCCTGGCCAAGTTCTTAGTTGCTGTGTATGTCTCGTGGTAG  1390

Qy   1321 GACTCTAGAACTGTAGAGGAAGGAAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAA  1380
          ||||| |||| |||||||        |||||||||||||| |||||||| || |||||||
Db   1391 GACTGTAGAA-------AAGGGAACTGAACATTCCAGAGCGTGTGAGTAGTGAATCACGTAAA  1442

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA  1411
          |||||||| ||||||||| |||| |||||||
Db   1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA  1473


RESULT 4
US-10-181-906-9
; Sequence 9, Application US/10181906
; Publication No. US20040053864A1
; GENERAL INFORMATION:
; APPLICANT: Karsenty, Gerard
; APPLICANT: Amling, Michael
; APPLICANT: Ducy, Patricia
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BONE FORMATION VIA
; TITLE OF INVENTION: MODULATION OF NEUROPEPTIDE Y ACTIVITY
; FILE REFERENCE: 9142-020-999
; CURRENT APPLICATION NUMBER: US/10/181,906
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/02040
; PRIOR FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-181-906-9

Query Match    55.1%;   Score 933.4;   DB 13;   Length 1679;
Best Local Similarity   81.8%;   Pred. No. 2.1e-236;
Matches 1154;   Conservative   0;   Mismatches 216;   Indels   41;   Gaps   5;

Qy      1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG   60
          ||||||||||||| || ||| ||||||||||||||||||||||||| ||||||||| |
Db    104 ATATACACTTCAGATAACTACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG  163

Qy     61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC  120
          ||||||| ||||||||||||||||||| ||||| |||||| |||||||||||||||||
Db    164 GAACCCTGTTTCCGTGAGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC  223

Qy    121 TTCATCATCTTCTTGACTCGGCAATGAGTCGGCAATGGTGTGATCCTGGTCATGGGTTAC  180
          ||||||||||| ||||| ||| ||||| |||||| ||| ||||||||||| ||||||||||
Db    224 TCCATCATCTTCTTAACTGGGCAATGATTGGGCAATGGTGTGATCCTGGTCATGGGGTTAC  283

Qy    181 CAGAGAGAGCTAAGGACGTAAGGACGATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  240
          |||||||||||||||||||||||||||||||||||||| ||||| |||||||||| ||||||||||
Db    284 CAGAGAGAAACTGAGAAGCATGACGGACAAGTACAGGCGGGTGAACCTGTCAGTGGCCGACCTC  343

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTTGATGCCAGTGACTGGTACTTTGGG  300
          ||||||||||||||||||||||||||||||||||| |||||||| |||||||||||||
Db    344 CTCTTTGTCATCACGCTCCCCTTCTGGGCAGTTGATTCCGCAAACTGGTACTTTGGG  403
```

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCTGCATA 1411
Db   1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1473

RESULT 5
US-10-342-887-912
; Sequence 912, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
;   APPLICANT: Dai, Hongyue
;   APPLICANT: He, Yudong
;   APPLICANT: Linsley, Peter S.
;   APPLICANT: Mao, Mao
;   APPLICANT: Roberts, Christopher J.
;   APPLICANT: Van 't Veer, Laura Johanna
;   APPLICANT: Van de Vijver, Marc J.
;   APPLICANT: Bernards, Rene
;   TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;   FILE REFERENCE: 9301-188-999
;   CURRENT APPLICATION NUMBER: US/10/342,887
;   CURRENT FILING DATE: 2003-01-15
;   PRIOR APPLICATION NUMBER: 60/298,918
;   PRIOR FILING DATE: 2001-06-18
;   PRIOR APPLICATION NUMBER: 60/380,710
;   PRIOR FILING DATE: 2002-05-14
;   PRIOR APPLICATION NUMBER: 10/172,118
;   PRIOR FILING DATE: 2002-06-14
;   NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-912

Query Match   55.1%;   Score 933.4; DB 13; Length 1679;
Best Local Similarity   81.8%;   Pred. No. 2.1e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTTCCAACAAG 60
          |||||||||||| ||||| ||| || ||||||||| ||||| |||||||||||||| |||
Db   104 ATATACACTTCAGATAATACACCGAGGAAAATGGGCTCAGGGGACTATGACTTCCATGAAG 163

Qy   61 GAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
          |||||||| ||||| ||||| |||| |||||||| | | |||||||||||||| |||
Db   164 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC 223

Qy   121 TTCATCAATCTTCTGACTGGCATAGTCGGCAATGTGGGCAATGGTCATCCTGGTCATGGGTTAC 180
          ||||| ||||||||| ||||||||| ||||||| ||||||| |||||||| ||||||||||||||
Db   224 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGTCATCCTGGTCATGGGTTAC 283

Qy   181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACGGCTGCACCTGTCAGTGGCTGACCTC 240
          |||||||||||| ||| |||||||||||||||||||| ||||||||||| ||||||||||
Db   284 CAGAAGAACTGAGAAGCATGACGACGAAGTACAGCGGCTGCACCTGTCAGTGGCGCACCTC 343

Qy   241 CTCTTGTCATCACACTCCCCTTCTCGGGCAGTTGATGCCATGCAGTCATGGCTGGTACTTTGGG 300
          |||||||||||| ||||| ||||| ||||||| |||||||||||||| || || ||||||||||
Db   344 CTCTTGTCATCACGCTTCCCTTCTCCGGCAGTTGATGCCATGCAGTCATGGCAAACTGGTACTTTGGG 403

Qy   301 AAATTTTTGTGTAAGGCGTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
          |||||||| |||||||| |||||||||||||| ||||| ||||||||||||||| |||||||
Db   404 AACTTCCTATGCAAGGCAGTCCATGTCATCTCATCCACACAGCAGTGTCCTC 463

Qy   361 ATCCTGGCCTTCATCAGCCTGACCGGGTACCTGCCCAGGACTCACTCGCCACCACCAGTCAA 420
          ||||||||||||||||| |||||||||||| ||||| ||||||| ||||||||||||| |||
Db   464 ATCCTGGCCTTCATCAGTCTGGCCGGGTACCTGCCCAGGACTCACTCGCCACCATCTAC 523

Qy   421 AGGCCAAGGAAAACTGCTGGCTCATGTCGGGTCTCGGTGGGACATCAGTCAGGGG 480
          |||||||||||||| ||||| |||||||||||| |||||||||||||||||||||
Db   524 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTTGGCGTCTGGATCCCCTGCCCCTC 583

Qy   481 CTCCTGACTATACCTGACTTCATCTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
          |||||||| |||| || ||| ||||| ||||| |||||||||||||||||||||
Db   584 CTGCTGACTATTCCCGACTTCATCTTGCCAACG 628

Qy   541 GATGACAGGTACATCGTGACCGCCTTACCCCGATAGCCTGTGATGGTGGTGTTTCAA 600
          |||| |||||||||||||||||| ||||||||| |||||||||| ||||||| |||||
Db   629 GATGACAGATATATCGTGACCGCCTTCTCACCCCAATGACTTGTGTGGGTGGTTGTTGTTCCAG 688

; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10/151,274
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/517,605
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-151-274-14

Query Match   55.1%;   Score 933.4;   DB 13;   Length 1679;
Best Local Similarity 81.8%;   Pred. No. 2.1e-235;
Matches 1154;   Conservative 0;   Mismatches 216;   Indels 41;   Gaps 5;

```
Qy     1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
           ||||||||||||||||| |||||||||||||| |||||| ||||||||||||||| |||||
Db   104 ATATACACTTCAGATAACTACTACACCGAGAAATGGCTCAGGGGACTATGACTCCATGAAG 163

Qy    61 GAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC 120
           |||||||||||||||||||||| ||||||||||  ||||||||||||||||||||||||||
Db   164 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCCACCATCTAC 223

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGGGTTAC 180
           ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   224 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGATTGGTCATCCTGGTCATGGGTTAC 283

Qy   181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGGCTGACCTC 240
           ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   284 CAGAAGAAACTGAGAAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGGCGACCTC 343

Qy   241 CTCTTTGTCATCACACTCCCCTTCGGCCAGTTGATGCCAGTTGATGATGCTGACTGGTACTTTGGG 300
           |||||||||||||||||||||| ||||||||||||||||||||||||||||| |||||||||||||
Db   344 CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCAGTTGATGATGCTGAAACTGGTACTTTGGG 403

Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATTTCCAGATATAAGAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   404 AACTTCCATGCAAGGCAGTCCATGTCATCTCACACTCACAGCAGTGTCCTC 463

Qy   361 ATCCTGGCCTTCATCAGCTGGACCGGTACCTCGCCATTGTCAACCTCTACAGCAGCGCGTTCTC 420
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   464 ATCCTGGCCTTCATCAGCTGGACCGGCACCGTCCGCCATCGTCACGCGCCACCAACAGTCAA 523

Qy   421 AGGCCAAGGAAACTGCTGGCGAAAGGCAGTCTATGTGGGCGTCTGAATCCCAGCCCTC 480
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   524 AGGCCAAGGAAGCTGTTGGCTGACAAAAGGTGGTCTATGTTGGCGTCTGATCCCTGCCCTC 583

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGAGTCAGCGTCAGCGGGGACACTCAGTCAGGGG 540
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   584 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG-----------TCAGTGAGGCA 628

Qy   541 GATGACAGGTACATCTGACGCCTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA 600
           |||||||||||||||| |||||||||| |||||||||||||||||||||||| ||||
Db   629 GATGACAGATATATCTGTGACCGGCTTCACCCCAATGACTTGTGGGTGGTGTTGTTCAG 688

Qy   601 TTCCAGCATATAATGGTGGTCTCATCCTCGCCGGCATCGTCATCCTCTCCTGTTACTGC 660
           ||||||||||||||||||| ||||||||||||||||||||| ||||||||||||||||||
Db   749 ATTATCATCTTCCAAGCTGTCACACTCGCACTCCAAGGGCCCAGGGAAGCGCAAGGACC 808

Qy   661 ATCATCATCTCTAAGCTGTCACACTGTCACTCCAAGAGCCCAGGCCTCAAGACG 720
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   749 ATTATCATCTTCCAAGCTGTCACACTCGCACTCCAAGGGCCCAGGGAAGCGCAAGGACC 808

Qy   721 ACAGTCATCTCATCCTAGCTTCTTGCCTGCTGCCTATTATGTGGGGATCAGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   809 ACAGTCATCCTCATCCCTGGCTTTTTTCGCCTGTTGGCTTCGCCTACTACACATTGGGATCAGC 868

Qy   781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
```

```
Db   584 CTGCTGACATTCCCGACTTCATCTTTGCCAACG-----------TCAGTGAGGCA 628

Qy   541 GATGACAGGTACATCTGACGCCTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAA 600
           |||||||||||||||| |||||||||| ||||||||||||||||||||||| ||||
Db   629 GATGACAGATATATCTGTGACCGGCTTCACCCCAATGACTTGTGGGTGGTGTTGTTCAG 688

Qy   601 TTCCAGCATATAATGGTGGTCTCATCCTCGCCGGCATCGTCATCCTCTCCTGTTACTGC 660
Db   749 ATTATCATCTTCCAAGCTGTCACACTCGCACTCCAAGGGCCCAGGGAAGCGCAAGGACC 808

Qy   721 ACAGTCATCCTCAGCTGTTGCTGTCTGA 748
Db   629 GATGACAGATATATCTGTGACCGGCTGCACCTCCAATGACTTGTGGGTGGTGTTGTTCAG 688

Qy   781 ACGTCATCCTCATCCTAGCTTCTTGCCTGCTGCCTATTATGTGGGGATCAGC 780
Db   809 ACAGTCATCCTCATCCCTGGCTTCTTGCCTGCTGCCTATTATGTGGGGATCAGC 868

Qy   781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
Db   869 ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGGTGTGAGTTTGAGAACACTGTG 928

Qy   841 CACAAGTGGATCCATCACAGAGAGCCCTGCCTTCTTCCACTGTGCCTGAACCCCATC 900
Db   929 CACAAGTGGATTCCATCACCGAGGCCCTAGCTTTCTTCCACTGTGTCTGAACCCCATC 988

Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTCAAAAGCTCGCCCAGCATGCACTCAACTCCATG 960
Db   989 CTCTATCTTTCCTTGGAGCCAAATTTAAAACCTCTGCCCAGCACCTCACCTTCTGTG 1048

Qy   961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
Db  1049 AGCAGAGGGTCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTT 1108

Qy  1021 TCCACCGGAGTCAGAATCCTCAGTTTCACTCCAGCTAACCCTTATGCAAAGACTTATAT 1080
Db  1109 TCCACTGAGTCGAGTCTTCAAGTTTCACTCCAGCTAA-------CACAGATGTAAAA 1160

Qy  1081 AATATATATATATATGATGAAGAACTTTTTTTTTTTTTTTTTTTACACATTTCCAGATATAAGAG 1140
Db  1161 GACTTTTTTTTATACGATAAATAAATAACTTTTTTTTTTTAAGTTACACATTTTTCAGATATAAAAG 1220

Qy  1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTTTTTTTGACTGTTGGGAGTTTATGTTCC 1200
Db  1221 ACTTTTTTTATACGATAAATATTGTTAATTGCTTGTTGGATTTTTGT------CTTGTGTTTC 1274

Qy  1201 TCTAGTTTTTGTGAGGTTGACTTAATTATATAAATATATTGTTTTTTGTTTTCATGT 1260
Db  1275 TTTAGTTTTTGTGAAGTTTAACTTAATTGACTTATTTATA---TAAATTTTTTTGTTTTCATAT 1330

Qy  1261 GAATGAGCGTCTAGGCAGGAGGAAGAAACTGAACATTCCAGAATGTGTGTAAATTGAATAAA 1320
Db  1331 TGATGTGTGTCTTAGGCAGGAGGAACCTGTGGCCAAGTTTCTTAGTTGCTGTATGTCTCGTGGTAG 1390

Qy  1321 GACTGTGAACTGTAGAGGAAGAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
Db  1391 GACTGTGAA--------AAGGGAACTGAACATTCCAGAACGTGTAGTAGCTGTAGTCACGTAAA 1442

Qy  1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
Db  1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1473
```

Query Match   55.1%;   Score 933.4;   DB 13;   Length 1679;
Best Local Similarity 81.8%;   Pred. No. 2.1e-236;
Matches 1154;   Conservative   0;   Mismatches 216;   Indels 41;   Gaps   5;

```
Qy     1 ATATACACTTCTGATAACTACTCTGAAGAAGTCGGGTCTGAGACTATGACTCCAACAG   60
         ||||||||||||| |||||   | ||||||||   | ||   ||||||||||  ||||||
Db   104 ATATACACTTCAGATAACTACACCGAGGAAATGGCTCAGGGACTATGACTCCATGAAG  163

Qy    61 GAACCCTGCTTCCGGGATGAAAAGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC  120
         |||||||||||| ||||||  | ||||| |||| ||  | ||||||||||||||||||
Db   164 GAACCCTGTTCCGTGAAGAAATGCTAATTTCAATAAAATCTTCCTGCCCACCATCTAC  223

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGTCATGGGTTAC  180
         | ||| |||| ||||||||||| || ||||||||| ||  ||||||||||||||||||
Db   224 TCCATCATCTTCTTAACTGGCATGTGGGCACAGTATCGGCAATGGATTCATCCTGTCATGGGTTAC  283

Qy   181 CAGAAGAAGCTAAGGAGCAGCATGACGGACAAGTACGGCTGCACCTGTCAGTGGCTGACCTC  240
         |||||||||  ||||||||||||||||||||| ||||||||||||||||||| || ||||
Db   284 CAGAAGAAGCTGAGGAGCATGACGGACAAGTACCAGGCTGCACCTGTCAGTGGCCGACCTC  343

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG  300
         |||||||||||||||| |||||  |||||||||||||||  ||| ||||| ||||||||
Db   344 CTCTTTGTCATCACAGCGTTCCCTTCTGGGCAGTTTGTGTCGCTGGCAAACTGGTACTTTGGG  403

Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATCATCTGTCCCATTGTCCAGCGCCACCAACAGTCAA  360
         |||| | ||| |||||| ||||||||||| ||||||||||||  |||||||||||||||||
Db   404 AACTTCCTATGCGAAGGCAGTCCATGTCATCACAGTCTCACACCTCTACAGCGAGTGTCCCTC  463

Qy   361 ATCCTGGCCTTCATCAGCTGACCGGTACCTCCGCCATTGTCCAGCGCCACCAACAGTCAA  420
         |||||||||||||||||||||||  |||| ||||||||||||||||||||||||||||||
Db   464 ATCCTGGCCTTCATCAGCTGACCGGTACCTCCGCCATCGTCCAGCGCCATCGTCCACCAACAGTCAG  523

Qy   421 AGGCCAAGAAACTGCTGGCAGTTCTATGTGGGCCTCTGAATCCCAGCCCTC  480
         ||||||||||||  || | ||| |||||| ||||||||| |||| ||||||
Db   524 AGGCCAAGGAGCTGTTGGCTGAAAAGGTGGTCATGTTGGCGTCTGAAAGTGCTCATGTTGGCTCTGGTAGTCCCCTC  583

Qy   481 CTCCTGACTATACCTGACTTCATCTTGCCCGATTCATCTTTGCCAACG----------  540
         |||| |||| |||||||||||||||||| |||||| |||||  ||||||
Db   584 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG------------TCAGTGAGGCA  628

Qy   541 GATGACAGGTACATCTGTGACCGCCTTACCCGATAGCCTGTGATGTGGTGTTTTCAA  600
         ||||||||||||||||| ||||||||||||||| || ||||| ||||||| |||||
Db   629 GATGACAGATATATCTGTGACCGCCTTCATCTGCCTGTATTGTCATCCTGTCTTGTGTGTTGTGTTCCAG  688

Qy   601 TTCCAGCATATAATGGTGGGTCTCATCGTCCCGGCATCGTCATCCTCCTGTTACTGC  660
         |||||||| ||||||| || |||| ||||||||| |||| || |||||||||||||
Db   689 TTTCAGCACATCATGGTTGGCCTTATCATCCTGCGTATTGTCATCCTGTCCTGCTGTATTGC  748

Qy   661 ATCATCATCCTAGCTCCAAGGGCCACCAGAAGGCCAAGGGCCTCAAGACG  720
         ||||||||| ||||||||||| |||||||||||||||||| |||||||||
Db   749 ATTATCATCTCCAAGGTCACTCACACTGTCACACTCCAAGAAGCGCCACCAGAAGGGCCTCAAGACC  808

Qy   721 ACAGTGCCTCCTCATCCTAGCTTCTTTGCTGGCTGGATCAGC  780
         ||||| ||||||||||||||| |||| |||||||| ||||||
Db   809 ACAGTGCCTCATCCTAGCTTCTTTGCTGGCTGCCTTACTACATTGGGATCAGC  868

Qy   781 ATCGACTCCTTCATCCTTTGGGATGCTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  840
         ||||||||||||||||| ||| || |||||||||||||||||| ||||||||||||||| |
Db   869 ATCGACTCCTTCATCCTCCTGGAATCATCAAGCAAGCAAGGGTGTGAGTTTGAGAACACTGTG  928

Qy   841 CACAAGTGGATCTCCATCACAGAGAGGCCCCATC  900
         ||| |||||||||||||||||||||||||||||||
Db   929 CACAAGTGGATTCCATCCATCACCACTCCTTCCTCCACTGTCCTTCGAACCCCCATC  988

Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGAGACTCAACTCCATG  960
         |||||||||| ||||| ||||||||||||||||||  ||||||||| |||||||||||
Db   989 CTCTATGCTTTCCTTGGAGCCAAGTTAAAAACCTCGCCCAGAGACGCACGCACTCACCTCTGTG  1048

Qy   961 AGCAGGAGGTCCAGCCTCAAGATCCTTTCCAAAGGAAAGGGGTGGACACTTCTTCCGTC  1020
         ||||||||| |||||||||||||| |||||| ||| | ||||||||||| |||||||||
Db  1049 AGCAGGAGGGTCAGCCTCAAGATCCTCTCAAGAAAAAGCGAAAAGGAAAGCGAGGTGGACATTCATCTGTT  1108

Qy  1021 TCCACCGAGTCAGAATCCTCCAGTTTTCACTCAGTTTTATGCAAAGACTTATAT  1080
```

```
Db   869 ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGCAAGGGTGTGAGTTTGAGAACACTGTG   928
Qy   841 CACAAGTGGATCTCCATCACAGAGAGGCCCTTCTTCCACCACTGTGTGAGAACCCCATC   900
Db   929 CACAAGTGGATTCCATCCATCACCACTCCTTCCTCCACTGTGTGAACCCCATC   988
Qy   901 CTCTATGCCTCCTCGGGGCCAAGTTCAAAAGCTCGCCCAGCATGCCCAGCACTCCATG   960
Db   989 CTCTATGCTTCCTTGGAGCCTCAAGATCCTTCAAGAGTGCTTCGCTCGTCAAAAAGCATCTGTT  1048
Qy   961 AGCAGGAGGTCCAGCCTCAAGATCCTTTCCAAAGGAAAGGGGTGGACACTTCTTCCGTC  1020
Db  1049 AGCAGGAGGGTCCAGCCTCAAGATCCTCAAGATCCTCAAAAGGAAAGCGAGGTGGACATTCATCTGTT  1108
Qy  1021 TCCACCGAGTCAGAATCCTCCAGTTTTCACTCAGTTTTATGCAAAGACTTATAT  1080
Db  1081 AATATATATATATATGAATAACAGAATAATAAATAATTTTTTTTTTCAGATATAAAGAG  1140
Qy  1081 AATATATATATATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAAAGAG  1140
Db  1161 GACTTTTTTTTATACAGATAAATAAACTTTTTTTTAAGTTACACATTTTCAGATATAAAAG  1220
Qy  1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC  1200
Db  1221 ACTGACCAAATATGTACAGTTTTGCTTGGTGTTTGT------CTTGTGTTTC  1274
Qy  1201 TCTAGTTTTTGTGAGGTTGACTTAATTATTATAAATATTGTTTTTTGTTTTCATGT  1260
Db  1275 TTTAGTTTTTGTGAAGTTTAATTGACTTATTATA----TAAATTTTTTTGTTTCATAT  1330
Qy  1261 GAATGACGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTAG  1320
Db  1331 TGATGTGTGTCTAGGCAGGACCTGTGGCCCAAGTTCTTAGTTGCTGTATGTCCGTGGTAG  1390
Qy  1321 GACTGTAGAACTGTAGAGGAAGAAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA  1380
Db  1391 GACTGTAGAA------AAGGAACTGAACATTCCAGAGGCGTGTAGTGAGTGAATCACGTAAA  1442
Qy  1381 GCTAGCCGTGACCTGCAGCTGTTGCTGCATA  1411
Db  1443 GCTAGAAAATGATCCCCAGCTGTTTATGCATA  1473
```

RESULT 7
US-10-172-118-912
; Sequence 912, Application US/10172118
; Publication No. US20030224374A1
GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003467
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-912

```
Db   224  TCCATCATCTTCTTAACTGGCATTGTGGGCAATGAGTGGTCATCCTGGTCATGGGTTAC  283

Qy   181  CAGAGAAGCTAAGGAGCAGCATGAGCGACGACGAACAAGCTACGGCTCACCTGTCAGTGGCTGACCTC  240
          |||||||||    ||||||||| |||||||||||||||||||| |||||||||||||||||||||
Db   284  CAGAGAAACTGAGAGAACATGAGCGACGACGAACAAGCTACAGCGGCTACCTGTCAGTGGCTGGCCGACCTC  343

Qy   241  CTCTTTGTCATCACACTCCCCTTCTCGGCGCAGTGATGCCATGGCCATGGCTGGTACTTTGGG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   344  CTCTTTGTCATCACGCTTCCCTTCTCGGCGCAGTGATGCCATGGCCAAACTGGTACTTTGGG  403

Qy   301  AAATTTTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTCTTC  360
          ||    ||||||||||||||||||  ||| ||||||| ||||| ||||||||||
Db   404  AACTTCCTATGCAAGGCAGTCCATGTCATCTCACAGTCAACCTCTACAGCAGTGTCCTC  463

Qy   361  ATCCTGGCCTTCATCAGCCTGACCGCTGGTACCTCGCCATTGTCCACGGCCACCAAACAGTCAA  420
          ||||||||||||||||||  |||  |||||||||   ||||||| |||||||||||||||||||
Db   464  ATCCTGGCCTTCATCAGTCTGACCGCGTCACCTGGCCATCGTCCACGCGCACCACCACAGTCAG  523

Qy   421  AGGCCAAGGAAACTGCTGGCTAGGAAGAAACCAGTCTAATGGGGCGTCTCGATCCCAGCCCTC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   524  AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTTGGGCGTCTGGATCCCTGCCCTC  583

Qy   481  CTCCTGACTATACCTGACTTCATCTTTGGCGACGTCAGCCAGGGGACTCAGTCAGTGGGG  540
          |||||||||    |||||||||||||    |||||||
Db   584  CTGCTGACTATTCCCGACTTCCCGACTTCTTTGCCAACG----TCAGTGAGGCA  628

Qy   541  GATGACAGTACAATCTGTGACCGCCTTACCCCGATAGCCTGGATGGTGGTGTTGTTGTTCAA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   629  GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTGTGGGTGGTGGTTGTTCAG  688

Qy   601  TTCCAGATATATGGTGGCTCTCATCCTGCCGCACGTCATCCTCTCCTGTTACTGC  660
          |||||||||||  ||||||  ||||||||||||||| |||||||||||||||||||
Db   689  TTTCAGCACATCATGGTGGCCTTATCCTGCTTATCCTGTATTCATCCTGCTATTGC  748

Qy   661  ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCTCAAGAGCGCAAGGCCCTCAAGACCG  720
          ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db   749  ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCTCAAGAGCGCAAGGCCCTCAAGACC  808

Qy   721  ACAGTCATCTCATCCTCAGCTTCTTGCCTGCCCATATATGTGGGATCAGC  780
          |||||||||||||| |||||||||||||||||||||||||  ||||||||
Db   809  ACAGTCATCTCATCCTCTGCTTCTTGCCTGTCCCGTTGTGCCTTACACATTGGGATCAGC  868

Qy   781  ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  840
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   869  ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG  928

Qy   841  CACAAGTGGATCCTCCATCACCAGAGGCCCTGAACCCATC  900
          ||||||||||||||||||||||||||||||  ||||||||
Db   929  CACAAGTGGATTTCATCACCGAGGCCCCTAGCTTTCTTCCACTGTTGTCTGAACCCCATC  988

Qy   901  CCTCTATGCCTTCCTCCGGGGCCAAGTCAAAAGCTCCGCCCAGTCACTCAACTCCATG  960
          ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db   989  CCTCTATGCTTTCCTTGGAGCCAAATTTAAAACCTCTGCCCAGCACGCCACTCACCTCTGTG  1048

Qy   961  AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1049  AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCCTGTT  1108

Qy  1021  TCCACGGAGTCAGAATCCTCCAGTTTCACTCAGCTAAACCCTTATGCAAAGACTTATAT  1080
          |||||||||||||||||  ||||||||||||||||||         ||||||||||
Db  1109  TCCACTGAGTCTGAGTCTTCAAGTTTCACTCAGCTAA------CACAGATGTAAAA  1160

Qy  1081  AATATATATATATATAATGAGGCTCCAACCCTTTTTATGTTACACATTTTCAGAATATAAGAG  1140
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  1161  GACTTTTTTTATACGATAAATAAACTTTTTTTTTAAGTTACACATTTTCAGATATAAAAAG  1220

Qy  1141  ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC  1200
          |||||||||||||||||||||||   |||||||||| ||||| |||||||||
Db  1221  ACTGACCAATATTGTACAGTTTTTTATTGCTGTGTGGGATTTTTGT------CTTGTGTTTC  1274

Qy  1201  TCTAGTTTTTGTGAGGTTTGACTTAATTATATAATAAATATTGTTTTTGTTTTTCATGT  1260
          |||||||||||||||||||||||||||    ||||||||||||||||||||||||||
Db  1275  TTTAGTTTTTGTGAAGTTTTAATTGACTTATTATA----TAAATTTTTTGTTTTTCATAT  1330
```

US-10-170-385-332

Query Match   55.1%; Score 933.4; DB 13; Length 1679;
Best Local Similarity 81.8%; Pred. No. 2.1e-236;
Matches 1154; Conservative   0; Mismatches 216; Indels   41; Gaps   5;

```
Qy     1  ATATACACTTCTGATAACTACTCCTGGAGAAGTGGGGTCTCGGAGACTATGACTCCAACAAG   60
          |||||||||||||||||||||||||||||| ||||||||||| |||||||||||||||||
Db   104  ATATACACTTCAGATAACTACTCACACCGAGGAAGTGGGGCTCAGGGGACTATGACTCCATGAAG  163

Qy    61  GAACCCTGCTTCCGGGATGAAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC  120
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   164  GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAAATCTTCCTGCCCACCATCTAC  223

Qy   121  TTCATCATCTTCTTGACTGGCATGGCCCAGTGTGGCAATGAGTGGTGATCCTGGTCATGGGTTAC  180
          || |||||||||||||||| |||||||  |||||||||||||  |||||||||||||||||||
Db   121  ...
```

Qy 481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGCAGGGGACATCAGTCAGGGG 540
         |||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db 584 CTGCTGACTATTCCCGACTTCATCTTTGCCCAACG------------------TCAGTGAGGCA 628

Qy 541 GATGACAGGTACATCTGTGACCGGCCTTTACCCGATAGCCTGTGATGGGGTGTTTCAA 600
         ||||| |||||| ||| |||||||||||||||| |||||||||||||| ||||||||||
Db 629 GATGACAGATATATCGTGACCGGCTCTACCCCAATGACTGTGGGGTGGTGTGTGTTCCAG 688

Qy 601 TTCCAGCCATATAATGGTGGTCTCCATCGTGCCCGCGGCATCGTCATCCTTCCTGTTACTGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 689 TTTCAGCACATCATGGTTGCCTTATCCTGCCTGCTGTATTGTCATCCTGTCTGCTATTGC 748

Qy 661 ATCATCATCCTCTAAGCTGCACCTGTCACACTCCAAGGGGCACCCAGCAGCCCTCAAGACG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 749 ATTATCATCCTCCAAGCTGCACACTCCAAGCTGCACACTCCAAGGCCCAAGGCCCTCAAGACC 808

Qy 721 ACAGTCATCCTCATCCTAGCTTCTTGTGCTGCTGGCTGCCCATTATGTGGGGATCAGC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 809 ACAGTCATCCTCATCCTGGCTTCTTCGCCTGTTCGCTGGCCGCTTACTACATTGGGGATCAGC 868

Qy 781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGACTTCGAGAGACATTGTG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 869 ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTGAGAACACTGTG 928

Qy 841 CACAAGTGGATCTCCATCACAGAGAGCCCTGCCTTCTTCCACTGTGCCTGACCCCCATC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 929 CACAAGTGGGATTTCCATCACCGAGCGCCCTAGCTTTCTTCCACTGTGTCTGAACCCCATC 988

Qy 901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGTCTGCGCCAGCATGCACTCAACTCCATG 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 989 CTCTATGCTTTCCTTGGAGCCAAATTAAAACCCTGCGCCAGCACGCACTCACCTCTGTG 1048

Qy 961 AGCAGAGGCTCCAGCCTCCAAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
         ||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db 1049 AGCAGAGGGTCCAGCCTCCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTT 1108

Qy 1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGATATATAT 1080
         |||||||||||| || |||||||||||||||| |
Db 1109 TCCACTGAGTCTGAGTCTTCAAGTTTCCACTCCAGCTAA-------CACAGATGTAAA 1160

Qy 1081 AATATATATATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATAAGAG 1140
         |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db 1161 GACTTTTTTTTATACGATAAATAAACTTTTTTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1220

Qy 1141 ACTGACCAGCTGTGTACAGTTGTACAGTTTTATTGACTGTGGGAGTTTATGTTCC 1200
         |||||||||||||||||||||||||||||||||||| |||||||||||
Db 1221 ACTGACCAATATTGTACAGTTTTTCAATAAAATCTTCCTGCCCACCATCTAC------CTTGTGTTTC 1274

Qy 1201 TCTAGTTTTTGTGAGGTTGACTTAATTATATAAATATTGTTTTTGTTTGTTCATGT 1260
         |||||||||||||||||||||||||||||||| |||||||| |||||||
Db 1275 TTTAGTTTTTGTGAAGTTTAATTGACTTATTTATA---TAAATTTTTTTTTTGTTCATAT 1330

Qy 1261 GAATGAGCGTCAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTATCGTGTGTAG 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1331 TGAGTGTGTCTAGGCAGGACCCGTGGCCAAGTTCTTAGTTGCTGTAGTGTCTCGTGGTAG 1390

Qy 1321 GACTGTAGAACTGAGAGGAGGAGAAACTGAACATTCCAGAAATTGAATAAA 1380
         |||||||| |||||||||||||||||||||||||||||||||
Db 1391 GACTGTAGAA-------AAGGGAACTGAACATTCCAGAGCGTGAGTGAATCACGTAAA 1442

Qy 1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
         |||||||||||||||||||||||||||||||
Db 1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1473

RESULT 9
US-10-225-567A-75
; Sequence 75, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;　APPLICANT: LifeSpan Biosciences
;　APPLICANT: Brown, Joseph P.
;　APPLICANT: Burmer, Glenna C.
;　APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 75
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-75

Query Match　　55.1%;　Score 933.4;　DB 15;　Length 1679;
Best Local Similarity　81.8%;　Pred. No. 2.1e-236;
Matches 1154;　Conservative　0;　Mismatches 216;　Indels　41;　Gaps　5;

Qy 1 ATATACACTTCTGATAACTACTCTGAGAAGTGGGTCTCGAGACCATGACTCCAACAAG 60
        ||||||||||||||||||||||| |||||||| |||||||||||||||||||||||
Db 104 ATATACACTTCAGATAACTAACTACACCGAGGAAATGGGCTCAGGGGACTCAGACTGA 163

Qy 61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
        |||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 164 GAACCCTGTTTCCGGTGAAGAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTAC 223

Qy 121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 224 TCCATCATCTTCTTAACTGGCATGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 283

Qy 181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTCACCGCTCGTCAGTGGCTGCACCTC 240
        ||||||||||||||||||||||||||||||||| ||||||| ||||||||||||
Db 284 CAGAAGAAACTGAGAAACTGGACATCAGCATGACGACAAGTACAGGCTGCACCTGGCCGACCTC 343

Qy 241 CTCTTTGTCATCACACTCCCCCTTCTGGGACAGTTGATGCCATGGCTGACTGGTTTGGG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 344 CTCTTTGTCATCATCACGCTTCCCCTTCTGGGCAGTTGATGTCCGTGGCAAACTGGTACTTTGGG 403

Qy 301 AAATTTTTGTGTAAGGCTGTCCATATCATCATCCTGCTCAACCTCTACACAGCAGCGGTTCTC 360
        ||||||| || ||||||||||||||||||||||||||||||||||||||||||||||||
Db 404 AACTTCCTATGCCAAGGCAGTCCATGTCCATGTCATCTCACAGTCAACCTCTACAAGCAGTGTCCTC 463

Qy 361 ATCCTGGCCTTCATCAGCCTGACCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAA 420
        |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db 464 ATCCTGCCTTCATCATCGGCCGCTACCGCGACGTCCACCACGCCACCAACAGTCAG 523

Qy 421 AGGCCAAGGAAACTGCTGGCTGAAAAGCAGCAGTCTATGTGGCGTCTGGATCCCAGCCCTC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 524 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTCATGTTGCGGTCTGGATCCCTGCCCCTC 583

```
;   APPLICANT: EOS Biotechnology, Inc.
;   TITLE OF INVENTION: No. US20030152926Ale1 Methods of Diagnosis of Angiogenesis,
;   TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
;   TITLE OF INVENTION: Modulators
;   FILE REFERENCE: 018501-000710US
;   CURRENT APPLICATION NUMBER: US/10/021,660
;   CURRENT FILING DATE: 2001-12-06
;   PRIOR APPLICATION NUMBER: US/09/784,356
;   PRIOR FILING DATE: 2001-02-14
;   PRIOR APPLICATION NUMBER: US 09/637,977
;   PRIOR FILING DATE: 2000-08-11
;   NUMBER OF SEQ ID NOS: 135
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 58
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-021-660-58

Query Match           55.1%;  Score 933.4;  DB 15;  Length 1679;
Best Local Similarity 81.8%;  Pred. No. 2.1e-236;
Matches 1154;  Conservative 0;  Mismatches 216;  Indels 41;  Gaps 5;

Qy       1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAG 60
           || ||| |||||| ||||||||| ||||||  | ||| ||| |||||||||  |||||||
Db     104 ATATACACTTCAGATAACTACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163

Qy      61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 120
           ||||||||||||| ||||||||||||||||||||||||| |||||||||||||||||||||
Db     164 GAACCCTGTTCCGGTGAGAGAAATGGTAATTCAATAAAATCTTCCTGCCCACCATCTAC 223

Qy     121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
           |||||||||||| ||| |||||||||||  |||||||| |||||||||||||||||||||
Db     224 TCCATCATCTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 283

Qy     181 CAGAAGAAGTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGCACCTC 240
           |||||||||| || ||| ||||||||||||||||||| |||||||||||||||||||||||
Db     284 CAGAAGAAACTGAGAGCATGACGGACAAGTACAGCGGCTGCACCTGTCAGTGGCCGCACCTC 343

Qy     241 CTCTTTGTCATCACACTCCCCTTCTGGCAGTTGATGCCATGGTGCTGACTGGTACTTTGGG 300
           ||||||||||||||||||||||||||||| ||||  | || ||||| |||||||||||||
Db     344 CTCTTTGTCATCACGCTTCCCTTCTGGCAGTTGATGCCGTGGCAAACTGTGACTTTGGG 403

Qy     301 AAATTTTGTGTGAAGGCTGTCCATATCATCTCAACCTCTACAGCAGCGCTTTTC 360
           ||||| | | | ||||||||||||||||||||||| |||||||||||| ||||||
Db     404 AACTTCCTATGCAAGGCAGTCCATGTCATCATCTCAACCTCTACAGCAGTGTCCCTC 463

Qy     361 ATCCTCGGCCTTCATCAGCCGGACCGGTACCTGCCATTGCCCACGGCCACCAACAGTCAA 420
           |||||||||| |||| |||||||  ||| |||||||| |||||||||||  ||||||||||
Db     464 ATCCTGCCTTCATCAGTCTGACCACGTTCCAGCTCCTGGCATCGTTCCAGGCCACCAACAGTCAG 523

Qy     421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCCAGTCTATGTGGCGTCTGGATCCCAGCCCTC 480
           |||||||||||||||||||||||||||||||||| |||||| ||||||||||| ||||||
Db     524 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGTCTATGTTGCGTCTGATCCCTGCCCTC 583

Qy     481 CTCCTGACTATACCTGACTTCATCTTTGCCCAGTCAGCCAGGGGGACATCAGTCAGGGG 540
           || ||||||| |||  || |||||| |||||||||||| |||||||||||||||||||||
Db     584 CTGCTGACTATTCCCGACTTCATCTTTGCCACG----------TCAGTGAGGCA 628

Qy     541 GATGACAGGTACATCTGTGACCGCCTTACCCCGATAGCCTGTGTGGATGGTGGTGTTCAA 600
           ||||| | ||| |||||||||||||| | |||| ||  ||||||||||| ||||||||||
Db     629 GATGACAGATATATCTGTGACCGCCTTATCCTGGTATTGTGTCATCGTCCTGCTATTGC 688

Qy     601 TTCCAGCATATAAATGGTGGGGTCTTCAATCATCACTCCTCCTGTACTGC 660
           ||||| ||| || || |||| |||| ||||| ||| ||||| ||  ||||
Db     689 TTTCAGCACATCATGGTTGGCCTTATCCTGCTGTAATTGCCTGTATTGC 748

Qy     661 ATCATCATCCTAAGCTGTCACACTCCAAGGCGCCACCAGAAGCGCAAGGCCCTCAAGACG 720
           || |||||| ||||||||||| |||||| |||||  |||||||||||||||| |||||
Db     749 ATTATCATCTCCAAGCTGTCACACTCCAAGGCGCCAGAAGCGCAAGGCCCTCAAGACC 808

Qy     721 ACAGTCATCATCCTCATCCTAGCTTTCTTTGCCTGCTGGCTGCCCATATTATGTGGGGATCAGC 780
           || |||||||||||||||||| |||||| || | |||| | ||  ||||||| | |||| |||||
```

```
Db     809 ACAGTCATCCATCCTGGCTTTCTTCGCCTGTTGGCTGCCTTACTACATTGGGATCAGC 868
           ||||||| |||||||||||||||||||||||||||||| ||||| |||||||||||||
Qy     781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAGCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
           ||||||||||||||||||||| |||||||||||||||||||||||||| |||| | ||||
Db     869 ATCGACTCCTTCATCCTCTGAAAATCATCAAGCAAGCAAGGGTGTGAGTTTGAGACACTGTG 928

Qy     841 CACAAGTGATCTCCATCACAGAGCGCCCTGCCCTGTTGCCTGACCCCATC 900
           |||||||| ||||| ||||||||||||||||||||  ||||||||||||||
Db     929 CACAAGTGAATTCCATCACCGAGCGCCCTGCCCTGTGTCTGCTGACCCCCATC 988

Qy     901 CTTTATGCCTTCCTCGGGCCAGTCCAAAGCTCTGCCCAGCATGCAGTCACTCAACTCCATG 960
           |||||||||||||||||||||||| ||||||||||||||| |||||| | || |||| |||
Db     989 CTCTATGCTTCCTTGGGAGCCAAATTTAAACCCTCGCCCAGGAAGGAAAGCGAGTGGACATTCACCTGTG 1048

Qy     961 AGCAGAGGCTCCAGCCTCAAGATCCTTCCAAAGAAAGCGGGTGGACACTCTTCCGTC 1020
           || ||||| ||||| ||||||||| ||||||||||| |||| |||||||  ||||
Db    1049 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCCAAAGGAAAGCGAGTGGACATTCATCTCGTT 1108

Qy    1021 TCCACGGACTCAGAATCCTCCAGTTTCACTCCAGCTAA------CACAGATGTAAAA 1080
           |||| ||| |||||||||||||||||| ||||||||| ||         |||||| |||
Db    1109 TCCACTGAGTCTGAGTCTTCAAGTTTTCACTCCAGCTAA 1140

Qy    1081 AATATATATATATATATATGATAAAGAACTTTTATGTTACACATTTCCAGATATAAGAG 1140
           ||||  |||||  |||||| |||| ||||  |||  ||||||  ||||||||||
Db    1161 GACTTTTTTTTATCAGATAAAATAACTTTTTTTTAAGTTACACATTTTCAGATATAAAG 1220

Qy    1141 ACTGACCAGTCTGTACAGTTTTTTTTTTTTTTTTAATTGACTGTGGGAGTTTATGTTCC 1200
           |||||||| ||| |||||| |||||  ||       |||| || | | ||||| ||||
Db    1221 ACTGACCAATATTGTACAGTTTTTTATTGCTTGTGTGGATTTTTGT------CTTGTGTTTC 1274

Qy    1201 TCTAGTTTTTTGTGAGGTTTGACTTAATTTATATAAATATTGTTTTGTTTCATGT 1260
           |||||||||||||||  |||||  ||| || | | |||       |||| |||||
Db    1275 TTTAGTTTTTTGTGAAGTTTTAATTGACTTATTTATA---TAAATTTTTTTTGTTTCATAT 1330

Qy    1261 GAATGAGAGCGTCTAGGCAGGACCGTGGCCAAGTTCTTAGTAGCGTGTTTATCTGTGTGTAG 1320
           ||| || |||||||||||| ||||||||||||||| ||||||| | || || | || |||||
Db    1331 TGATGTGTGTCTAGGCGAGGACCGTGGGCCAAGTTCTTAGTTCTTAGTTGCTGTATGTCTCGTGGTAG 1390

Qy    1321 GACTGTGAAACTGTAGAGGAAGAAACTGAACATTCCAGAAATTGTGGTGTAAATTGAATAAA 1380
           ||||| ||||                  ||||||| ||||| || |||  || |||||
Db    1391 GACTGTAGAA--------AAGGGAACTGAACATTCCAGAGCGGTGTAGTGAATCACGTAAA 1442

Qy    1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
           ||| ||| || |||||| || || ||| |||
Db    1443 GCTAGAAATGATCCCCAGCTGTTTATGCATA 1473
```

```
RESULT 11
US-10-341-434-177
;  Sequence 177, Application US/10341434
;  Publication No. US20030215835A1
;  GENERAL INFORMATION:
;   APPLICANT: OriGene Technologies
;   TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
;   FILE REFERENCE: 9U 204 205 R1
;   CURRENT APPLICATION NUMBER: US/10/341,434
;   CURRENT FILING DATE:  2003-07-18
;   PRIOR APPLICATION NUMBER: US 60/348,164
;   PRIOR FILING DATE: 2002-01-15
;   PRIOR APPLICATION NUMBER: US 60/348,119
;   PRIOR FILING DATE: 2002-01-15
;   NUMBER OF SEQ ID NOS: 238
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 177
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (89)..(1144)
;   OTHER INFORMATION:
US-10-341-434-177
```

Db 1109 TCCACTGAGTCTGAGTCTTCAAGTTTTCACTCCAGCTAA------ CACAGATGTAAAA 1160
Qy 1081 AATATATATATATATATGTAAGAACTTTTTTTATGTTACACATTTTCCAGATATAAGAG 1140
Db 1161 GACTTTTTTTTATCGATAATAAAATAACTTTTTTTTAAGTTACACATTTTCAGATATAAAAG 1220
Qy 1141 ACTGACCAGTCTTGTACAGTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC 1200
Db 1221 ACTGACCAATATTGTACAGTTTTATTTGTTGTTGGATTTTGT------ CTTGTGTTTC 1274
Qy 1201 TCTAGTTTTTGTGAGGTTGACTTAATTATATAAAAATATTGTTTTTTGTTTCATGT 1260
Db 1275 TTTAGTTTTTGTGAAGTTTAATTGACTTATTTATA------ TAAATTTTTTTGTTTCATAT 1330
Qy 1261 GAATGAGCGTCAGGCAGGACCTGTGCCAAGTTCTTAGTAGCTGTTTATCGTGTTGGTAG 1320
Db 1331 TGATGTGTGTCTAGGCAGGACCTGTGCCAAGTTCTTAGTTGCTGTATGTCTCGTGGTAG 1390
Qy 1321 GACTGTGTAGAACTGTAGAGGAGGAAGAAACTGAACATTCCAGAAATGTGTGAATAAA 1411
Db 1391 GACTGTAGAA------ AAGGGACTGAACATTCCAGAGACGTGTGTAGTGAATCACGTAAA 1442
Qy 1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA 1411
Db 1443 GCTAGAAATGATCCCCCAGCTGTTTATGCATA 1473

RESULT 12
US-09-971-392-20
; Sequence 20, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Template ID: 284616.2
; NAME/KEY: unsure
; LOCATION: 1706
; OTHER INFORMATION: a, t, c, g, or other
US-09-971-392-20

Query Match 55.1%; Score 933.4; DB 10; Length 1711;
Best Local Similarity 81.8%; Pred. No. 2.2e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy   1 ATATACACTTCTGATAACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCACACAAG 60
Db 139 ATATACACTTCAGATAACTTACACCGAGAGAAATGGGCTCAGGGGACTATGACTCCATGAAG 198
Qy   61 GAACCCTGCTTCCGGGATGAAAAGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
Db 199 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCTGCCCACCATCTAC 258
Qy 121 TTCATCATCTTCTTGACTGCATAGTCGGCAATGGTGATCCTGGTCATGGGTTAC 180
Db 259 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGTCATCCTGGTCATGGGTTAC 318
Qy 181 CAGAGAAGCTAAGGACATGACGGACAAGTACCGGCTCACCTGTGCACTGTGACCTC 240

```
Db   1366  TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTCTATGTCTCGTGGTAG  1425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321  GACTGTAGAACTGTAGAGGAAGAGAAACTGACACATTCCAGAGATGTGGTAAATTGAATAAA  1380

Db   1426  GACTGTAGA-------AAGGGAACTGAACATTCCAGAGCGTGTAGTGAATCACGTAAA  1477
           |||||||||          |||  ||  ||||    || ||||
Qy   1381  GCTAGCCGTGATCCTCAGCTGTTGCTGCATA  1411

Db   1478  GCTAGAAATGATCCCCAGCTGTTTATGCATA  1508
           |||||   ||||| ||||||||  ||||||||
```

```
RESULT 13
US-10-101-510-459
; Sequence 459, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 459
;   LENGTH: 1711
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1706)
;   OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-459

Query Match      55.1%;  Score 933.4;  DB 15;  Length 1711;
Best Local Similarity  81.8%;  Pred. No. 2.2e-236;
Matches 1154;  Conservative  0;  Mismatches 216;  Indels  41;  Gaps  5;
```

```
Qy     1  ATATACACTTCGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139  ATATACACTTCAGATAACTAACACCAGAGAAATGGGCTCAGGGACTATGACTCCATGAAG  198

Qy    61  GAACCCTGCTTCCGGATGAAAACGTCATTCAATAGGATCTTCCTGCCACCATCTAC  120
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   199  GAACCCGTTTCCGTGAAGAAAATGCTAATTCAATAAAATCTTCCTGCCACCATCTAC  258

Qy   121  TTCATCATCTTCTTGACTGCGCATAGTCGGCAATGATGGTTGGTGATCCTCATGGTTAC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259  TCCATCATCTTCTTAACTGGCATTGTGGGCAATGATGGTCATCCTGGTCATGGGTTAC  318

Qy   181  CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCAGTGGCTGGTGGTACTTGGG  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319  CAGAAGAAACTGAGAAGCATGACGAGCAAGTACACAGGCTGCAGTGCCGGCGACCCTC  378

Qy   241  CTCTTTGTCATCACACTCCCCTTCGGGGCAGTGATGCGGCAGTGGTGACTTTGGG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   379  CTCTTTGTCATCACGCCTCCCTTCTGGGGCAGTGATGCCGTGGCAAACTGGTACTTTGGG  438

Qy   301  AAATTTTGTGTAAGGCTGTCCATATCATTTTCCAGATATAAGAG  360
          |||||||||||||||||||||||||||||||||||||||
Db   439  AACTTCCATATGCAAGGCAGTCCATGTCATCTCACAGTCATCACAGTCAGCAGTGTCCTC  498

Qy   361  ATCCTGGCCTTCATCAGCCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGAATCCCAGCCCTC  420
          |||||||||||||||||||||||||||||||||||||||||
Db   499  ATCCTGGCCTTCATCAGCCTGGACGCTGGAACCGCTACCTGCCGGCCCACCAACAGTCAG  558

Qy   421  AGGCCAAGGAAGGAAGCTGTTGCGTGAAAAGGTGGTCTATGTTGGCCGTCTGGCGT  480
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   559  AGGCCAAGGAAGGAAGCTGTTGCGTGAAAAGGTGGTCTATGTTGGCCGTCTGGCGT  618

Qy   481  CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTGTCAGGGG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   319  CAGAAGAAACTGAGAAGCATGACGGACAAGCAAGATACAGGCTGCACCTGTCAGTGGCCGACCTC  378
          ||||||||

Qy   241  CTCTTTGTCATCACACTCCCCTTCGGGGCAGTGATGCCATGGCCTGACTCGTACTTTGGG  300
          |||||||||

Db   379  CTCTTTGTCATCACGCCTCCCTTCTGGGCAGTGATGCCGTGGCAAACTGGTACTTGGG  438

Qy   301  AAATTTTGTGTAAGGCTGTCCATATCATCTCACAGTCATCACAGCAGCGTTCTC  360

Db   439  AACTTCCATATGCAAGGCAGTCCATGTCATCTCCACAGTCATCACAGCAGCGTCCTC  498

Qy   361  ATCCTGGCCTTCATCAGCGTGACCGGTACCTCGCCATTGTCACCGCCACCAACAGTCAA  420

Db   499  ATCCTGGCCTTCATCAGCTGGACCGCTACCTGCCGGCCCACCAACAGTCAG  558

Qy   421  AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCCTGAATCCCAGCCCTC  480

Db   559  AGGCCAAGGAAGGAAGCTGTTGCGTGAAAAGGTGGTCTATGTTGGCCGTCTGATGCCTC  618

Qy   481  CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTGTCAGGGG  540
```

CORRESPONDENCE ADDRESS:
  ADDRESSEE: Genentech, Inc.
  STREET: 1 DNA Way
  CITY: South San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94080
COMPUTER READABLE FORM:
  MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/104,063
  FILING DATE: 24-June-1998
  CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/701265
  FILING DATE: 22-AUG-1996
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/664228
  FILING DATE: 06-JUN-1996
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/076093
  FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 07/810782
  FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
  NAME: Love, Richard B.
  REGISTRATION NUMBER: 34,659
  REFERENCE/DOCKET NUMBER: P0706P2C2
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415/225-5530
  TELEFAX: 415/952-9881
  TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
  LENGTH: 1737 base pairs
  TYPE: Nucleic Acid
  STRANDEDNESS: Single
  TOPOLOGY: Linear

US-09-104-063-3

Query Match   55.1%;   Score 933.4; DB 9;   Length 1737;
Best Local Similarity 81.8%; Pred. No. 2.2e-236;
Matches 1154; Conservative   0; Mismatches 216; Indels 41; Gaps   5;

Qy   1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGGACTATGACTCCAACAAG 60
            |||||||||||| |||||||||||||| |||||| || ||||||||||||||||||||
Db   106 ATATACACTTCAGATAACTAACACCGAGGAAACGTCATCTCGGGGACTATGACTCATGAAG 165

Qy   61 GAACCCTGCTTCCCGGGATGAAAACGTCCATTTCAATAGGGATCTTCCTGCCACCATCTAC 120
            |||||||||||||||| |||| | ||||||| | ||||||||||| |||||||||||||
Db   166 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCCTTCCTGCCCACCATCTAC 225

Qy   121 TTCATCATCTTCTGACTGGCATAGTCGGCAAATGGATTGGTGATCCTGGTCATGGGTTAC 180
            |||||||||||||||||||| |||||||||||||||||||||| |||| |||||||||||
Db   226 TCCATCATCTTCTTAACTGGCATTGTGGGCAAATGGATTGGTGATCATCCTGGTCATGGGTTAC 285

Qy   181 CAGAAGAAGCTAAGGAGCATGGAGGACAAGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
            ||||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||
Db   286 CAGAGAAACTGAGAAGCATGACGAGCATGACGACACAGTACAGCAAGTGCCACCTGTCAGTGGGCCGACCTC 345

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGCAGTGGCAGTGCGCTGATGCCATGGGTGGTACTTTGG 300
            |||||||||||||||||||| |||||||||||||| ||||||||||||||| ||||||||||||
Db   346 CTCTTTGTCATCACGCTTCCCTTCTCGGGCAGTTGGCAGTTGATGCCGTGGCAAACTGGTACTTTGG 405

Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTCTACACAGCAGTGTCCTC 360
            ||||||||||| |||||||||||||||| ||||||||| |||||||||||
Db   406 AACTTCCCATGCAAGGCAGTCCATGTCATCTCATCTACACAGCAGTGTCCTC 465

Qy   361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCCCCATTGTCCACGCCACCAACAGTCAA 420

Db   619   TCAGTGAGGCA 663
              |||||| |||| |||||
Qy   541 GATGACAGGTACATCGTGACCGCCCTTACCCCAGCTGTGATGGTGGTGTTCAA 600
            |||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   664 GATGACAGACAGTATATCTGTGACCGCGTTCTACCCCCATGACTGTGGGGTGGTGTGTTCCAG 723

Qy   601 TTCCAGCATATAAATGGTGGGTCTCATCCTGTCATCCTCCTGTGTACTGC 660
            ||||| |||||| ||||||||||||||||||||||||||||||||||||
Db   724 TTTCAGCACATCATGGTTGGCCTTATCGTGTTGTCATCCTGTCCTGCTATTGC 783

Qy   661 ATCATCATCTCTAAGCTGTCACACTCCAAGAGCCACCAGAAGCCAAGGCCCTCAAGACG 720
            |||||||||||||||||||||| ||||||||||||| ||||||||||||||||||||||
Db   784 ATTATCATCTCCAAGCTGTCACACTCCAAGAGCCGCCACCGAAGCCCTCAAGACC 843

Qy   721 ACAGTCATCCTCATCCTAGCTTCTTCTTGCCTGCTGCTGCCATATTATGTGGGGGATCAGC 780
            ||||| |||||| || | ||||||||||| ||||| |||| ||| || ||||||||||||
Db   844 ACAGTCATCCTCATCCTGGCTTCTTCTGCCTGCTCGTGTTGCTTACTACACATTGGGGATCAGC 903

Qy   781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGACATTGTG 840
            |||||||||| ||||||||||||||| |||||||||||||||| ||||||| |||||||||
Db   904 ATCGACCCCTTCATCCTCCTGGAAATCATCAGCAAGGGGTGGTGAGTTTGAGAACACTGTG 963

Qy   841 CACAAGTGGGATCTCCATCACAGAGAGGCCCTGCCTTCTTCCACTGTGCCTGAACCCCATC 900
            ||||||||||||| |||||||||||| |||||| |||||||| || |||||||||||||||
Db   964 CACAAGTGGATTTCCATCACCGGAGGGCCCTAGCTTTCTTCCTGCCTGTGCCTGAACCCCATC 1023

Qy   901 CTCTATGCCTTCCTGGGGCCAAGTCAAAAGCTCGCCCAGATGCACTCAACTCCATG 960
            |||||||| |||||||||||| |||||||||| |||| |||||||| ||||||||||
Db   1024 CTCTATGCTTTCCTTGGAGCCAAATTTAAAAACCTCTGCCAGCCCAGCACTCACCTCTGTG 1083

Qy   961 AGCAGAGGCGTCCAGCCTCAAGATCCTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
            ||||||||| |||||||||||||||||| |||||||||||| |||||||||||| |||||
Db   1084 AGCAGAGGGTCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGGACATTCATCTGTT 1143

Qy   1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCCAGCTACCCTTATGCAAAGACTTATAT 1080
            |||| |||||| ||||||||||||||||||| ||||| || |||| ||||||||||
Db   1144 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAA------CACAGATGTAAAA 1195

Qy   1081 AATATATATATATATATACGATAAATAACTTTTTTAAGTTACACATTTTCAGATATAAAAG 1140
            ||||||||||||| |||| ||| ||||||||||||||||||||||||||||||||
Db   1196 GACTTTTTTTTATAACAGTTTTTTTTTT ||||| TTTAAGTTACACATTTTCAGATATAAAG 1255

Qy   1141 ACTGACCAGTCTTGTACAGTTTTTTTTTT TTAATTGACTGTTGGGAGTTATGTTCC 1200
            |||| ||||||||||||||| ||| ||| | |||| |||||||| |||||
Db   1256 ACTGACCAATATGTACAGTTTTATTGCTTGTTGGATTTTGT------CTTGTGTTTC 1309

Qy   1201 TCTAGTTTTGTGAGGTTGACTTAATTATTATATAAATATGTTTTTTGTTTCATGT 1260
            ||| |||||||||| ||||| |||||||||| | ||||||||||||||||||
Db   1310 TTTAGTTTTTGTGAGTTTAATTGACTTATTTATA---TAAATTTTTTTTGTTTCATAT 1365

Qy   1261 GAATGAGGCGTCTAGGCAGCAGAAACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTAG 1320
            |||| |||||| ||||||||||||||||||||||||||||||||| ||||||| ||||||||
Db   1366 TGATGTGTGTCAGGCAGGAGCACCTGTGGCCAAGTTCTTAGTAGCTGTGTATGCTGTGGTAG 1425

Qy   1321 GACTGTAGAACTGTAGAGGAAGAAAACTGAACATTCCAGAATGTGTGGTAAATTGAATAAA 1380
            |||||||||| ||||||| |||||||||||||||||||||||||||||||||||
Db   1426 GACTGTAGAA------AAGGGAACTGAACATTCCAGAGCGTGTAGTGAGTGTGAATCACGTAAA 1477

Qy   1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCGCATA 1411
            ||||| |||||| |||||||| ||||||
Db   1478 GCTAGAAATGATCCCCCAGCTTGTTTATGCATA 1508

RESULT 14
US-09-104-063-3
; Sequence 3, Application US/09104063
; Patent No. US2002016856A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PF4A Receptors
; NUMBER OF SEQUENCES: 6

US-10-666-689-3
; Sequence 3, Application US/10666689
; Publication No. US20040037830A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Human PF4A Receptors, Nucleic Acid Encoding and
; TITLE OF INVENTION: Antibodies Binding Thereto
; FILE REFERENCE: P0706P22D2C1
; CURRENT APPLICATION NUMBER: US/10/666,689
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 09/104,063
; PRIOR FILING DATE: 1988-06-24
; PRIOR APPLICATION NUMBER: US 08/701,265
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/664,228
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/076,093
; PRIOR FILING DATE: 1993-06-11
; PRIOR APPLICATION NUMBER: US 07/810,782
; PRIOR FILING DATE: 1991-12-19
; PRIOR APPLICATION NUMBER: US 07/677,211
; PRIOR FILING DATE: 1991-03-29
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
;   LENGTH: 1737
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-666-689-3

Query Match   55.1%;   Score 933.4;   DB 13;   Length 1737;
Best Local Similarity   81.8%;   Pred. No. 2.2e-236;
Matches 1154;   Conservative   0;   Mismatches 216;   Indels 41;   Gaps 5;

Qy   1 ATATACACTTCTGATAACTACTTCTGAAGAAGTGGGGTCTGAGAGACTATGACTCCAACAAG 60
           ||||||||||||||||||||||||||||||||||||||| || |||||||||||||||||
Db   106 ATATACACTTCAGATAACTACACACCGAGAAAATGGGCTCAGGGGACTATGACTCCATGAAG 165

Qy   61 GAACCCTGCTTCCGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC 120
           |||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   166 GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATAAAATTTCCTGCCACCATCTAC 225

Qy   121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATGGTTGGTCATCCTGTCATGGGTTAC 180
           ||||||||||| ||| || |||||||| || |||||||||||||||||||||||||||||
Db   226 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGATGGTTGGTCATCCTGTCATGGGTTAC 285

Qy   181 CAGAAGAAGCTAAGGAGCATGACAGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 CAGAAGAAACTGAGAAGGATGACAGGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC 345

Qy   241 CTCTTGTCATCACACTCCCCTTCTGGGCAGTGATGGCAGTGATCTTTGGG 300
           ||||||||||||||||| |||||||||||| ||||||||||||||||||||
Db   346 CTCTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCAAACGTGGTACTTTGGG 405

Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATCCACTGTCAACCTCACAGCAGCGTCTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 AACTTCCTATGCAAGGCCAGTCCATGTCATGTACCAGTCAACCTCACAGCAGTGTCTC 465

Qy   361 ATCCTGGCCTTCATCAGCCTGACCGGTTGCGACTCTTGCCATTGTCCACCAACAGTCAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466 ATCCGGCCTTCATCAGCTCATCGACCTGGCCATCGGCCACCGCCACCAACAGTCAG 525

Qy   421 AGGCAAGGAAACTGCTGGCTGAAAGCAGTCTATGTGGGGCTCTGGGGCGTCTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 AGGCCAAGGAAGCTGTTGGCTGGTCTATGTTGGGCTGGCTCTGCCCTGCCCTC 585

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
           |||||||||||||||||||| || |||||||||||||||
Db   586 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG------------TCAGTGAGGCA 630

Qy   541 GATGACAGGACACATCTGTGACCGCCTTACCTGTGAGCGATGTTGGTGTGTTTCAA 600
           || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   631 GATGACAGATATATCGTGACCGGCTTCTACCCCAATGACTTGTGGGTGGTGTGTTCCAG 690

```
Qy    601  TTCCAGCATATAATGTGGGTCTCATCCTGCCGGCGATCGTCATCCTCTCGTGTACTGC  660
           || ||| ||||||| |||||||||||||||||| |||||||| |||||| |||||| |||
Db    691  TTTCAGCACATCATGGTTGGCCTTATCCTGCCTGTATTGTCATCCTGTCCTGCTATTGC  750

Qy    661  ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCCGCAAGGCCCTCAAGACG  720
           |||||| |||||||||||| ||||||||||||| |||||||||||| |||||||||||||||
Db    751  ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACC  810

Qy    721  ACAGTCATCCTCATCCTAGCTTCTTGCCTGCTGCCATATTATGTGGGGATCAGC  780
           |||||||||||||||||||  ||||||| |||||||||||| ||| ||||||||||||||||
Db    811  ACAGTCATCCTCATCCTGGCTTCTTGCCTGCTGCCTTACTACATTGGGGATCAGC  870

Qy    781  ATCGACTCCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  840
           ||| ||||||| ||| |||| |||||||||||||| ||||||||| |||||||||||||||
Db    871  ATCGACTCCCTTCATCCTCCTGGAAATCATCAAGCAAGGGGTGTGAGTTTGAGAACACTGTG  930

Qy    841  CACAAGTGGATCTCCATCACAGAGAGGCCCTGCCTTCTTCCACTGTTGCCTGAACCCCATC  900
           ||||||||| |||| |||||||||||| |||| |||| |||||||||| |||||||||||||
Db    931  CACAAGTGGATTTCCATCACCGAGGGCCCTAGCTTTCTTCCACTGTTGTCTGAACCCCATC  990

Qy    901  CTCTATGCCTTCCTCGGGGCCAAGTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG  960
           |||||| ||||| |||| ||| ||||| ||  ||| || ||| || ||||| ||||| | |
Db    991  CTCTATGCTTTCCTTGGAGCCAAATTAAAACCTCTGCCCAGCACGCACTCACCTCTGTG  1050

Qy    961  AGCAGAGGCCCAGCCTCAAGATCCTTTCCAAAGGAAAAGCGGGTGGACACTCTTCCGTC  1020
           ||| || ||| |||||| |||||| |||| |||||| |||||| ||||| | |||| |||
Db    1051 AGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTT  1110

Qy    1021 TCCACGGAGTCAGAATCCTCCAGTTTCACTCAGCCTAACCCTTATGCAAAGACTTATAT  1080
           |||| | |||||| ||||||||||||||| ||||||||           ||||| |||||
Db    1111 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCAGCCTAA--------CACAGATGTAAAA  1162

Qy    1081 AATATATATATATATATGATAAAGAACTTTTTATGTACACATTTCCAGATATAAGAG  1140
           ||     ||||||||| ||||| |||| ||||||||| ||||||||| |||||||| |
Db    1163 GACTTTTTTTTATACGATAAATAACTTTTTTTAAGTTACACATTTTCAGATATAAAAG  1222

Qy    1141 ACTGACCAGTCTGTACACAGTTTTTTTTTTTTTTTTTAATTGACTGTTGGGAGTTTATGTTCC  1200
           ||||||||| |||| ||| ||||||||||||| |||||||| || |||||| ||||| |||
Db    1223 ACTGACCAAATATGTACACAGTTTTTTTTATTGCTTGTGTTGGATTTTTGT------CTTGTGTTTC  1276

Qy    1201 TCTAGTTTTTTGTGAGGTTGACTTAATTTATATAAATATATTGTTTTTGTTTTCATGT  1260
           | ||| |||| |||||| |||| ||| ||| ||| |||||||     |||| |||| |||
Db    1277 TTTAGTTTTTTGTGAAGTTAATTGACTTATTTATA----TAAATTTTTTTGTTTCATA  1332

Qy    1261 GAATGAGGCGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG  1320
           ||| || | ||| |||| || ||||| ||||||| ||||| |||||| ||| |||||| |||
Db    1333 TGATGTGTGTCTAGGCAGGACCTGTGGCCAAGTTCTTAGTTGCTGTATGTCTCGTGGTAG  1392

Qy    1321 GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCCAGAAATGTGTGGTAAATTGAATAAA  1380
           |||||||||     |||| ||| |||| ||||||| |||||| ||||| |||||||||||||
Db    1393 GACTGTAGAA-------AAGGGAACTGAACATTCCAGAGCGGTGTAGTGAATCACGTAAA  1444

Qy    1381 GCTAGCCGTGATCCTCAGCTGTTGCTGCATA  1411
           ||||||  |||| ||||||| |||||
Db    1445 GCTAGAAATGATCCCCAGCTGTTTATGCATA  1475
```

```
            GenCore version 5.1.6
     Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     May 17, 2004, 05:50:31 ; Search time 4889.52 Seconds
                                     (without alignments)
                          10345.919 Million cell updates/sec

Title:          US-09-367-052-7
Perfect score:  1694
Sequence:       1 atatacacttctgataacta.........taaaagtcaatgataaactt 1694

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27513289 seqs, 1493109027 residues

Total number of hits satisfying chosen parameters:      55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      EST:*
           1:  em_estba:*
           2:  em_esthum:*
           3:  em_estin:*
           4:  em_estmu:*
           5:  em_estov:*
           6:  em_estpl:*
           7:  em_estro:*
           8:  em_htc:*
           9:  gb_est1:*
          10:  gb_est2:*
          11:  gb_htc:*
          12:  gb_est3:*
          13:  gb_est4:*
          14:  gb_est5:*
          15:  em_estfun:*
          16:  em_estom:*
          17:  em_gss_hum:*
          18:  em_gss_inv:*
          19:  em_gss_pln:*
          20:  em_gss_vrt:*
          21:  em_gss_fun:*
          22:  em_gss_mam:*
          23:  em_gss_mus:*
          24:  em_gss_pro:*
          25:  em_gss_rod:*
          26:  em_gss_phg:*
          27:  em_gss_vrl:*
          28:  gb_gss1:*
          29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	709.4	41.9	1091	13	BX397685	BX397685
2	694.8	41.0	1201	13	BX396365	BX396365
3	692	40.9	1201	13	BX421287	BX421287
4	688.6	40.6	1201	13	BX421511	BX421511

```
 5    684    40.4   906  13   BU523025   BU523025   AGENCOURT
 6   681.4   40.2  1051  13   BX358346   BX358346   BX358346
 7   672.2   39.7  1051  12   BM545259   BM545259   AGENCOURT
 8   666.6   39.4  1201  13   BX353253   BX353253   BX353253
 9   647.6   38.2  1201  13   BX462627   BX462627   BX462627
10   624.6   36.9  1201  13   BX418530   BX418530   BX418530
11   621.6   36.7   793   9   AA182270   AA182270   mt83f01.r
12   621.2   36.7   639   9   AA816049   AA816049   vr14h02.r
13   608.2   35.9  1201  13   BX462182   BX462182   vr14h02.r
14    603    35.6   952  13   BX362477   BX362477   BX362477
15   602.2   35.5   872  14   CD251331   CD251331   AGENCOURT
16   582.4   34.4   945  13   BX443258   BX443258   BX443258
17   580.2   34.3   911  13   BX443334   BX443334   BX443334
18   575.8   34.0   904  10   BF100790   BF100790   601753725
19   575.8   34.0   910  12   BI762229   BI762229   603049139
20    572    33.8   600  13   BU919690   BU919690   603-58 M
21    569    33.6   891  12   BI824663   BI824663   603033815
22    564    33.3   886  13   BX443228   BX443228   BX443228
23   558.6   33.0   950  13   BQ718617   BQ718617   AGENCOURT
24   552.6   32.6   765  12   BM387369   BM387369   UI-R-CN1-
25   551.6   32.6   582  12   BG145042   BG145042   ut75a01.y
26   549.6   32.4   749  14   CD466324   CD466324   LeukoN2_3
27   549     32.4  1099  13   BX397684   BX397684   ut75a01.y
28   547.6   32.3   581  14   CF540844   CF540844   UI-M-GW0-
29   546.8   32.3   928  14   CD389217   CD389217   AGENCOURT
30   544.4   32.1   958  12   BG174412   BG174412   60234232
31   539.6   31.9  1034  12   BM920800   BM920800   AGENCOURT
32   539.4   31.8  1201  13   BX358340   BX358340   BX358340
33    533    31.5   876  14   CA454255   CA454255   AGENCOURT
34   532.4   31.4  1201  13   BX421141   BX421141   BX421141
35   529.2   31.2   827  13   BI761664   BI761664   603046395
36   524.8   31.0   528  12   BG277031   BG277031   uv10e03.y
37   524.8   31.0   750  12   BG613352   BG613352   60261234
38    522    30.8   955  12   BG173867   BG173867   602333928
39   520.8   30.7   753  12   BG915636   BG915636   602814106
40   510.6   30.1   585  10   BE627479   BE627479   uu52b07.y
41   510.2   30.1   875  12   BI756157   BI756157   603030061
42    507    29.9  1201  13   BX418529   BX418529   BX418529
43   502.4   29.7   976  13   BU146456   BU146456   AGENCOURT
44   491.4   29.0   913  12   BM051973   BM051973   603638877
45   490.4   28.9   782  14   CB985366   CB985366   AGENCOURT
```

                          ALIGNMENTS

```
RESULT 1
BX397685
LOCUS       BX397685                1091 bp    mRNA    linear   EST 13-MAY-2003
DEFINITION  BX397685 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI041YO13 5-PRIME, mRNA sequence.
ACCESSION   BX397685
VERSION     BX397685.1  GI:30621359
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1091)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI041AH07QP1&cluster=1272.r. Contact :
            Feng Liang Email: fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
```

```
Db    902  GCATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTG  961

Qy    839  TGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGAACCCCA  898
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    962  TGCACAAGTGGATTCCATCACCGAGGCCGCCCTAGCTTYTTCCACTGTTGTTGACCCCA  1021

Qy    899  TCCTCTCATGCCTTCCTCGGGGCCAAGTCAAAAGTCTGCCCAGCATGCACTCAACTCCA   958
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1022  TCCTCTCATGCCTTCCTTGGGAGCCAAATTTAAAACCTGCCCAGCACGCACTCACCTCTG  1081

Qy    959  TGAGCAGAGG  968
               ||||||||||
Db   1082  TGAGCAGAGG  1091


RESULT 2

BX396365
LOCUS       BX396365                1201 bp    mRNA      linear    EST 13-MAY-2003
DEFINITION  BX396365 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI015YL20 5-PRIME, mRNA sequence.
ACCESSION   BX396365
VERSION     BX396365.1    GI:30609085
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 1272.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI015DF10QP1&cluster=1272.r.  Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI015DF10QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI015YL20"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer.Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match      41.0%;  Score 694.8;  DB 13;  Length 1201;
Best Local Similarity  83.9%;  Pred. No. 5.9e-104;
Matches  842;  Conservative   7;  Mismatches  134;  Indels  21;  Gaps   5;

Qy      1  ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG   60
               ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    134  ATATACACTTCAGATAACTACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG  193

Qy     61  GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCACCATCTAC  120
               ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    194  GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTAC  253

Qy    121  TTCATCATCTCTTGACTGGCATAGTCGGCAATGATTGGTCGGCAATGGATCGGTCATGGGTTAC  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254  TCCATCATCTTCTTAACTGGCATTGGGCATTGGTGGTCATCGGTCATGTCATGGGTTAC  313
```

COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DG004AD05QP1&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DG004AD05QP1.

FEATURES　　　Location/Qualifiers
source　　　1..1201
　　　　　　/organism="Homo sapiens"
　　　　　　/mol_type="mRNA"
　　　　　　/db_xref="taxon:9606"
　　　　　　/clone="CS0DG004YG09"
　　　　　　/tissue_type="B CELLS (RAMOS CELL LINE)"
　　　　　　/cell_line="RAMOS CELL LINE"
　　　　　　/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
　　　　　　/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
　　　　　　with a NotI-oligo(dT) primer. Five prime end enriched,
　　　　　　double-strand cDNA was digested with Not I and cloned into
　　　　　　the Not I and EcoRV sites of the pCMVSPORT 6 vector.
　　　　　　Library was not normalized."

ORIGIN

Query Match　　40.9%;　Score 692;　DB 13;　Length 1201;
Best Local Similarity　76.6%;　Pred. No. 1.7e-103;
Matches 847;　Conservative 43;　Mismatches 193;　Indels 23;　Gaps 4;

Qy　　　1 ATATACACTTCTGATAACTACTCTGAAGAGAGTGGGTCTGGAGACTATGACTCCAACAAG 60
　　　　　|||||||||||||||| |||| | |||||||||||||||||||||||||||||
Db　　 118 ATATACACTTCAGATAACTACACCGAGGAAATGGCTCAGGGACTATGACTACTGAAG 177

Qy　　　61 GAACCCTGCTTCCGGGATGAAAACGTCCATTCAATAGGATCTTCCTGCCCACCATCTAC 120
　　　　　|||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db　　 178 GAACCCTGTTTCCGTGAAGAAAATCTAATTCAATAAAATCTTCCTGCCCACCATCTAC 237

Qy　　 121 TTCATCATCTCTTGACTGGCATAGTCGGCAATGATTGTGATCCTGGTCATGGGTTAC 180
　　　　　|||| |||| |||||||||||||||||||||||||| |||||||||||||||
Db　　 238 TCCATCATCTCTTAACTGGCATTGTGGGCAATGATTGGTCATCCTGGTCATGGGTTAC 297

Qy　　 181 CAGAAGAAGCTAAGGAGCATGACGGACAAGAGCTGGACCACCGTGTCAGTGGCTGACCTC 240
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 298 CAGAAGAAACTGAGAGCATGACGGACAAGAGCTGGACCACCGTGTCAGTGGCCGACCTC 357

Qy　　 241 CTCTTTGTCATCACACTCCCCTTCTGGGGCATCGTCATCCTTCTCCTGTTACTGC 300
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 358 CTCTTTGTCATCACGCTTCCCTTCTGGGGCAGTTGATGCCATGGCCGTGGCAAACTGTACTTTGGG 417

Qy　　 301 AAATTTTGTGTAAGGCTGTCATCTCATATCATCTCACACGCCACATTGTCCACGCCACCAACAGTCAA 360
　　　　　|||||||||||||||||||| || ||| ||||||||||||||||||||||||||
Db　　 418 AACTTCCTATGCAAGGCAGTCCATGTCATCATCTCACACGTCACCAGTCCACAGTGTCCTC 477

Qy　　 361 ATCCTGGCCTTCATCAGCCTGGACCGGTACCTCACGCCATTGTCCACGCCATCGTCCTCC 420
　　　　　||||||||||||||||||||||||||||||||||||||
Db　　 478 ATCCTGGCCTTCATCATCAGTCTGGACCGGTACCTCACGCCATCGTCCTCGTCCTGCCCTC 537

Qy　　 421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCAATGTGGGCGGTCTGGATCCCAGCCCTC 480
　　　　　|||||||||| ||| |||||||||||||||||||||||||||||||||||||||
Db　　 538 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCATGTTGGCGTCATGTGGGACATCAGTCAGGGG 597

Qy　　 481 CTCCTGACTATACCTGACTTCATCTTTGCCACGTCAGCCAGGGGGACATCAGTCAGGGG 540
　　　　　|||||||||||||||||||||||||||||||||||||||||
Db　　 598 CTGCTGACTATTCCCGACTTCATCTTCTTTGCCAACG------------TCAGTAAGGCA 642

Qy　　 541 GATGACAGGTACATCTGTGACCGCCTTACCCCGATAGCCTGTGATGGTGGTGTTTCAA 600
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　 643 GATGACAGATATACTGTGACCGCCTTCACCCCAATGACTTGTGGGTGGTGTGTTCCAG 702

Qy　　 601 TTCCAGCATATAATGTGGGTCATCATCCTCCGCCCGGCATGTCATCCTCCTGTTACTGC 660

Qy　　 181 CAGAAGAAGCTAAGGAGCATGACGGACAAGAGCTGGACCACCGTGTCAGTGGCTGACCTC 240
　　　　　||||||||||||||||||||||||||||||||||||
Db　　 314 CAGAAGAAACTGAGAGCATGACGGACAAGAGCTGGACCACCGTGTCAGTGGCCGACCTC 373

Qy　　 241 CTCTTTGTCATCACACTCCCCTTCTGGGGCAGTTGATGCCATGGCCGTGACTGACTTTGGG 300
　　　　　||||||||||||||||||||||||||||||||||
Db　　 374 CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCGTGGCAAACTGTACTTTGGG 433

Qy　　 301 AAATTTTGTGTAAGGCTGTCCATATCATCTCAACCTCTACACTGTCCACGCAGCGGTTCTC 360
　　　　　|||||||||||||| ||| ||| ||| ||||||||||||
Db　　 434 AACTTCCTATGCAAGGCAGTCCATGTCATCATCTCACACAGTGTCCTC 493

Qy　　 361 ATCCTGGCCTTCATCAGCCTGACCGGTACCTGCCACCGTCCAGCCCCTC 420
　　　　　||||||||||||||||||||||||||||||||||||
Db　　 494 ATCCTGGCCTTCATCATCAGTCTGACCGGTACCTGCCACCGTCCAGCCCCTC 553

Qy　　 421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCCTATGTGGGCGTCTGGATCCCTGCCCTC 480
　　　　　|||||||||| ||| |||||||||||||||||||||||||||
Db　　 554 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCATTGTTGGCGTCTGGATCCCTGCCCTC 613

Qy　　 481 CTCCTGACTATACCTGACTTCATCTTTGCCACGTCAGCCAGGGGGACATCAGTCAGGGG 540
　　　　　|||||||||||||||||||||||||||||||||||||||
Db　　 614 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG------------TCAGTAGGGCA 658

Qy　　 541 GATGACAGGTACAATCGTGACCGGCTTTACCCCGATAGCCTGTGATGGTGGTGTTTCAA 600
　　　　　|||||||||| |||||||||||||||||||||||||||||
Db　　 659 GATGACAGATATACTGTGACCGCCTTCTCACCCCAATGACTTGTGGGTGGTGTGTTCCAG 718

Qy　　 601 TTCCAGCATATAAATGTGGGTCGTCATCCTCGCCCGGCATCGTCATCCTTCTCCTGTTACTGC 660
　　　　　||||||||| ||| || || |||||||||||||||||||
Db　　 719 TTTCAGCCACCATCATGGTTGCCTTATCCTGCCTGGTATTGTCATCCTGCCTGCCTATTGC 778

Qy　　 661 ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCAAGGAGCGCAAGGCCCTCAAGACG 720
　　　　　|||||||||||| ||| ||||||||||||||||||||||||||||||
Db　　 779 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAAGGAGCGCAAGGCCCTCAAGACC 838

Qy　　 721 ACAGTCATCCTCCATCCTAGCTTTCTTTGCCTGCTGGCCATATATGTGGGATCAGC 780
　　　　　||||||||||||||||||| |||||||||||||||||||||
Db　　 839 ACAGTCATCCTCCATCCTCATCCTGGCTTTCTTTGCCTGCTGGCTACTCATTGGGATCAGC 898

Qy　　 781 ATCGACTCCTTCATCCTTTGGAGTCATCAAGCAAGGAGTGACTTGACTTCAGAGACATTGTG 840
　　　　　||||||||||||||||||| ||||||||||||||||||||
Db　　 899 ATCGACTCCTTCATCCTTCCTGGAAATCATCATCAAGGAGTGCAAGGGTGTGAGTTGAGAAACACTGTG 958

Qy　　 841 CACAGTGGATCCCATCACAGAGAGCCCTCAAGCCCTCTCCTGTTGCCTGAACCCCATC 900
　　　　　||||||||| ||| |||||||||||||||||||||||||||||||||
Db　　 959 CACAGTGGAT-TYCATCACCGAGAGCCCTTCCTTCTTCTTCCAGTGTGTCTGAA-CCCATC 1016

Qy　　 901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCCACTCAACTCCATG 960
　　　　　||||||||||||||||||||||||||||| || |||||||||||||
Db　　 1017 CTCTATGCTTTCCTTGGRGCAAATTTAAAAACTCT--GCCARCACGCACTCACCTCTGTT 1073

Qy　　 961 AGCAGAGGCTCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGG 1004
　　　　　||||||||| || || ||||||||||||||||||||
Db　　 1074 AGCAGAGGGKCMWAG-CYCAAGAGTATCTCTCMAARGGAAGCCGRGGGG 1116

RESULT 3
BX421287
LOCUS　　BX421287　　　　1201 bp　　mRNA　　linear　　EST 15-MAY-2003
DEFINITION　BX421287 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
　　　　　clone CS0DG004YG09 5-PRIME, mRNA sequence.
ACCESSION　BX421287
VERSION　BX421287.1　GI:30766068
KEYWORDS　EST.
SOURCE　　Homo sapiens (human)
ORGANISM　Homo sapiens
　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　1　(bases 1 to 1201)
AUTHORS　Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE　　Full-length cDNA libraries and normalization
JOURNAL　Unpublished (2001)

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match   40.6%;   Score 688.6; DB 13; Length 1201;
Best Local Similarity   84.8%;   Pred. No. 6.1e-103;
Matches 792; Mismatches 121; Conservative 6; Indels 15; Gaps   2;

```
Qy     1 ATATACACTTCGATAACTACTCTGAGAAGTGGGTCTGGAGACTATGACTCCAACAAG   60
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   142 ATATACACTTCAGATAACTACTACMCCGAGAAATGGCTCAGGGACTATGACTCCATGAAG  201

Qy    61 GAACCCTGCTTCCGGATGAAAACGTCATTCATTGCCTGAACCCATCTAC   120
           |||||||||||||||||||||||||||||||||||||||||||||:||
Db   202 GAACCCTGTTCCGTGAAGAAAATGCTAATTCAATAAAMTCTTCCTGCCACCATCTAC  261

Qy   121 TTCATCATCTTCTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGTGTCATGGGTTAC  180
           |||||||:|||||||||||||||||||| ||||||:|||||| ||||||||||||||
Db   262 TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTTCATGGGTTAC  321

Qy   181 CAGAAGAAGCTAAGGAGCATGACGGACAAGAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   322 CAGAAGAAMCTGAGAAGCATGACGACGACAAGGCTGCACCTGTCAGTGGCCGACCTC  381

Qy   241 CTCTTTGTCATCACACTCCCTTCTGGGCAGTTGATGCAGTGCATGCTGGTACTTTGGG  300
           |||||||:||||||||||||||||||:|||||||:|||||||||||||||||||||||
Db   382 CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCGTGCAAACTGGTACTTTGGG  441

Qy   301 AAATTTTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGGTTCTC  360
           |||||:|:|| || ||||||||||||||||||:||||||||||||||:||||:|||||
Db   442 AACTTCCTATGCMAGGCAGTCATGTCATCTACACACAGTCAACCTCTACAGCMCGTGTCCTC  501

Qy   361 ATCCTGGCCTTCATCAGCCTGACCGGTACCTGCCATTGTCCACGCCACCAACCAGTCAA  420
           ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||:|
Db   502 ATCCTGGCCTTCATCAGTCTGACCATGTCTGGACCGCTACCTGCCACCMCCACCAGTCAG  561

Qy   421 AGGCCAAGAAACTGCTGCTGAAAAAGCAGTCTATGTGGGGCCTGATCCCCCAGCCCCTCC  480
           ||||||||||||||||||| ||||||||||||:|||||||||||| ||||||:||||||
Db   562 AGGCCAAGAAGAAGTGTTGCTGAAAAAGGTGGTCTATGTTGGCCGTCTGGTCATCCTGCCCCTC  621

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCGACGTCAGCCAGGGGACATCAGTCAGTCAGGGG  540
           |||||||||||||||| |||||||||||:|||||||||||||||||||||||||
Db   622 CTGCTGACTATTCCCGACTTCATCTTTGCCAACGTCAGTCAGTCAGGGGCA----------  667

Qy   541 GATGACAGTACATCTGTGACCGGCTTTACCCGATAGCCTGTGGATGGTGGTGTTCAA  600
           |||||||||||||||||||||||||||||||||||||:||||||||:|||||:||||
Db   668 GATGACAGATATATCTGTGACCGGCTTCACACCCCAATGACTTGTGGGGTGGTTGTGTTCCAG  727

Qy   601 TTCCAGCATAATAGGTGGTCTCATCCTCCGGGATCACTGCATCCTCCTGTTACTGC  660
           |||||||:|||||||||||||||||||||||||||||||:|||||:|||||||||||
Db   728 TTTCAGCACATCATGGTTGGCCTTATCATCCTTCGTATTGTCATCCTGCCTGCTATTGC  787

Qy   661 ATCATCATCTCAAGCTGTCACACTCCAAGGGCCTCAAGAGCCTCAAGACG  720
           |||||||||||||||||||||| ||||||||||||||||||| ||||||:||
Db   788 ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCAAGGGCCCCCTCAAGACC  847

Qy   721 ACAGTCATCCTGCCTGGCTTCTTCTTGCTCGTGCTGGCTGCCATATTATGTGGGGATCAGC  780
           ||||||||||||||| |||||||||||||||:|||||||||||||||||:||||||||||
Db   848 ACAGTCATCCTGCCTTGGCTGCTCGTGCTGGCCTGCCTTACTACATTGGGATCAGC  907

Qy   781 ATCGACTCCTTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  840
           ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||:||
Db   908 ATCGACTCCTTCATCCTTCCTGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG  967

Qy   841 CACAAGTGATCTCCATCACAGAGCCCTCGCCTTCTTCCTGTTGCCTGAACCCATC  900
           ||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db   968 CACAAGTGGATTTCCATCACCGAGGCCCTAGCTTCTTCCCACTGTGTCTGAA-CCCATC 1026

Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTTCAAAGGT  934
           |||||:|| |||| || |||| ||| ||| ||
```

```
Db   1027 CTCTATGCTTCYTTGGAGCAAATTAAAACTCT 1060

RESULT 5
BU523025
LOCUS       BU523025              906 bp    mRNA    linear   EST 13-SEP-2002
DEFINITION  AGENCOURT 10154335 NCI_CGAP_Co24 Mus musculus cDNA clone
            IMAGE:6529663 5', mRNA sequence.
ACCESSION   BU523025
VERSION     BU523025.1  GI:22833475
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14130 row: c  column: 07
            High quality sequence stop: 668.
FEATURES             Location/Qualifiers
     source          1..906
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6529663"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Co24"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Query Match        40.4%;  Score 684;  DB 13;  Length 906;
Best Local Similarity  98.1%;  Pred. No. 3.6e-102;
Matches 714;  Conservative  0;  Mismatches 10;  Indels  4;  Gaps  2;

Qy    1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAAG 143

Qy   61 GAACCCTGCTTCCGGGATGAAAAGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  144 GAACCCTGCTTCCGGGATGAAAAGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 203

Qy  121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGGGTTAC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  204 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATTGGTGATCCTGGTCATGGGTTAC 263

Qy  181 CAGAGAAGCTAAGGAGCATGACGGACAGAAGTACCCGGCTGCACCTGTCAGTGGCTGACCTC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  264 CAGAGAAGCTAAGGAGCATGACGGACAGAAGTACCCGGCTGCACCTGTCAGTGGCTGACCTC 323

Qy  241 CTCTTTGTCATCACACTCCCCTTTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  324 CTCTTTGTCATCACACTCCCCTTTGGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 383

Qy  301 AAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  384 AAATTTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 443
```

```
Db   361 ATCCTGGCCTCATCAGCCTGGACCGGTACCTGCCATGTCCACGGCCACCAACAGTCAA 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   444 ATCCTGGCCTCATCAGCCTGGACCGGTACCTGCCATGTCCACGGCCACCAACAGTCAA 503

Qy   421 AGGCCAAGGAAACTGCTGCTGAAAAGGCAGTCATGTGGCGGTCTGGAGTCCCAGCCCTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   504 AGGCCAAGGAAACTGCTGCTGAAAAGGCAGTCATGTGGCGGTCTGGAGTCCCAGCCCTC 563

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGAGTCAGTCAGTCAGGGG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   564 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGAGTCAGTCAGTCAGGGG 623

Qy   541 GATGACAGGTACATCTGTGACCGCCTTACCCGATAGCCTGTGGATGGTGGTGTTTCAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   624 GATGACAGGTACATCTGTGACCGCCTTACCCGATAGCCTGTGGATGGTGGTGTTTCAA 683

Qy   601 TTCCAGCATATAATGGT-GGGTCTCATCCTGCCCGGCATCGTCATCCTCTCGTTACTG 659
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   684 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCTCGTTACTG 743

Qy   660 CATCATCATCTCTAAGCTGTCACACTCCAAGG---CCACCAGAAGCGGCAAGGCCCTCAA 716
         |||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db   744 CATCATCATCTCTAAGCTGTCACACTCCAAGGGGCCACCAGAAGCGGCAAGGCCCTCAA 803

Qy   717 GACGACAG 724
         ||||||||
Db   804 GACGACAG 811

RESULT 6
BX358346
LOCUS       BX358346             1201 bp    mRNA    linear   EST 05-MAY-2003
DEFINITION  BX358346 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CSODI036YJ17 5-PRIME, mRNA sequence.
ACCESSION   BX358346
VERSION     BX358346.1  GI:30370215
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fullength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI036CE09QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI036YJ17"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with NotI and cloned into the NotI and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match        40.2%;  Score 681.4;  DB 13;  Length 1201;
Best Local Similarity  85.3%;  Pred. No. 9.1e-102;
Matches 783;  Conservative  4;  Mismatches 116;  Indels 15;  Gaps  2;

Qy    1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAACAAG 60
        ||||||||||||||| ||||||  ||| || || |||||||||||||||||||||||
```

```
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 1051)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue procurement: Invitrogen
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM12720 row: e  column: 12
           High quality sequence start: 20
           High quality sequence stop: 753.
FEATURES           Location/Qualifiers
    source         1..1051
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:5726963"
                   /tissue_type="hippocampus"
                   /lab_host="DH10B"
                   /clone_lib="NIH_MGC_124"
                   /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
                   (destroyed); Site_2: NotI; RNA source male hippocampus,
                   age 27. Library is oligo-dT primed and directionally
                   cloned (EcoRV site is destroyed upon cloning). Average
                   insert size 1.4 kb, insert size range 0.9-4 kb. Library is
                   normalized and enriched for full-length clones and was
                   constructed by C. Gruber (Invitrogen). Research Genetics
                   tracking code 012."
ORIGIN

Query Match            39.7%;  Score 672.2; DB 12;  Length 1051;
Best Local Similarity  83.9%;  Pred. No. 3e-100;
Matches  825;   Conservative   0;  Mismatches 138;  Indels  20;  Gaps  5;

Qy      1 ATATACACTTCGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60
            |||||||||||||||||||||||||||||||||||||| | |||||||||||||||||
Db     83 ATATACACTTCAGATAACTACCACCGAGGAAATGGGCTCCAGGGGACTATGACTCCATGAAG 142

Qy     61 GAACCCTGCTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC 120
            |||||||||| |||||| ||||||||||   |||||| ||||||||||| ||||||||||
Db    143 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCCACCATCTAC 202

Qy    121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGATGGTGATCCTGGTCATGGGTTAC 180
            |||||||||| |||| ||| |||||||||||||| |||||| | ||||||||||||||
Db    203 TCCATCATCTTCTTAACTGGGCATTGTGGGCAATGGATTGGTCATCCTGGTCATGGGTTAC 262

Qy    181 CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTCACCTGTCAGTGGCTGACCTC 240
            |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||
Db    263 CAGAAGAAACTGAGAAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCCGACCTC 322

Qy    241 CTCTTTGTCATCACACTCCCCTTCTGGCCAGTTGATGCCTGACTGGTACTTTGGG 300
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    323 CTCTTTGTCATCACACGCCTCCCCTTCTGGCCAGTTGATGCCTGCAAACTGGTACTTTGGG 382

Qy    301 AAATTTTTGTGTAAGGCTGTCCATATCATCCACAGCAGCCGTTCTC 360
            |||| ||||||||||||||||||||||||||||||||||||||||
Db    383 AACTTCCTATGCAAGGCAGTCCATGTCATCCATGTCAACAGTCAACCTCTACACAGCAGTCTCTC 442

Qy    361 ATCCTGGCCTTCATCAGCCTGGACCGTGACCCGGTACCTCGCCATTGTCCAGGCCACCAGAAGACG 420
            |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    443 ATCCTGGCCTTCATCAGTCTGGACCGTGACCCGGTACCTCGCCATCGTCCAGGCCACCAGAAGACG 502

Qy    421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGAATCCCAGCCCTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 AGGCCAAGGAAGGAGCTGTTGGCTGAAAAGGTGGTCTATGTGGGCGTCTGGAATCCCAGCCCTC 562
```

/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match   39.4%;   Score 666.6;   DB 13;   Length 1201;
Best Local Similarity 82.6%;   Pred. No. 2.4e-99;
Matches 808;   Conservative 6;   Mismatches 146;   Indels 18;   Gaps 4;

```
Qy     1 ATATACACTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAACAG  60
           ||||||||||||||||||| ||||||||||||||||||| |||||| ||||||||||
Db   151 ATATACACTTCAGATAACTACACCGAGGAAATGGCTCAGGGACATAGACTCCATGAAG 210

Qy    61 GAACCCTGCTTCCGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCACCATCTAC 120
           |||||||||||||||||||  || |||||||||||||| ||||| |||||||||||||
Db   211 GAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCCTGCCACCATCTAC 270

Qy   121 TTCATCATCTTCTTGACTGCATAGTCGGCAATGATTGTGTGATCCTGGTCATGGGTTAC 180
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   271 TCCATCATCTTCTTAACTGCGCATTGTGGGCAATGATTGGTCATCCTGGTCATGGGTAC 330

Qy   181 CAGAGAAGCTAAGGAGCATGACGGACAAGCACCGGCTGCACCTTCAGTGGCTGACCTC 240
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   331 CAGAGAAACTGAGAGAGCATGACGGACAAGCACTGCACGGCTGCACCTTCAGTGGCCGACCTC 390

Qy   241 CTCTTTGTCATCACACTCCCCTTCCTTCTGGGCAGTTGATGGCCATGGCCATGGTACTTTGGG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   391 CTCTTTGTCATCACAGTCTTCTGGGGCAGTTGATGGGCAGTTGATGCTGGGCAAACTGGTACTTTGGG 450

Qy   301 AAATTTTGGTAGGCTGCCATATCATCATCATCTACACTGTCAACCTCTACAGCAGGGTTCTC 360
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   451 AACTTCCATGCAAGGCAGTCATGTCATCATCATCTACACAGTCAACCTCTACAGCAGTGTCCTC 510

Qy   361 ATCCTGGCCTTCATCAGCCGTACCTCGGACCGGCCATTGTCCACGCCACCAACAGTCAA 420
           |||||||||||||||||||||||||||||||||||||||:|  ::||||||||:||
Db   511 ATCCTGGCCTTCATCAGTCGGACCGGCCATCGTCCACGCCACCAACAGTTAG 570

Qy   421 AGGCCAAGGAAAACTGCTGGCTGAAAAGGCAGTTATGTGGGCGTTCTGAGTCCCAGCCCTC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   571 AGGCCAAGGAAGCTGTTGGCTGGAAAAGGTGGTCTCATGGTGGCGTCTGGATCCCTGCCCTC 630

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
           |||||||||||||||||||||||||||||||||||----------|||:||||
Db   631 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG----------TCAGTANGGCA 675

Qy   541 GATGACAGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTCAA 600
           ||||||| |||||||||||||||||||||||||||||||:| |||||||||||||||
Db   676 GATGACAGATATATCTGTGACCGCCTTCTACCCCAATGACTTGTGGTGTGGTGTTCCAG 735

Qy   601 TTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCATCGTCATCCTCCTGTTACTGC 660
           |||||||||||||||| ||||||| ||||||||||||||||||||||||||||||||
Db   736 TTTCAGCACATCATGGTTGGCCTTATCCTGCCTGTGTGATTGTCATCCTGCCTGCTATTGC 795

Qy   661 ATCATCATCCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCCCTCAAGACG 720
           |||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   796 ATTATCATCCAACCTGTCACACTCCAAGGGCCACCAGAAGCGCCCCAAGACC 855

Qy   721 ACAGTCATCATCCTCATCCTAGCTTTCTTTGCCGGCTGGGCTGCCATATTATGTGGGGATCAGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   856 ACAGTCATCCTCATCCTCGGCTTCTTCGCCTGGCTGTTKGGGTGCCTGTTACTACATTGGGATCAGC 915

Qy   781 ATCGACTCCTCATCCTTTGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   916 ATCGACTCCTCATCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 975

Qy   841 CACAAGTGGATCCCATCACAGAGGCCCTTCGCCTTCTCACTGTTGCCTGAACCCCATC 900
           ||||||||||||||||||||||||||||||||||||| ||||||| ||||:||
Db   976 CACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTTCTT-CACTGTTGCTGAA-CCCATC 1033

Qy   901 CTCTATGCCTTCCTCGGGGCCAAGTCAAAAGTCTGCCCAGCATGCACTCAACTCCATG 960
           |||||||:| |||||||||||||||||:|:|||||||||||||||||||||||||
Db  1034 CTCTATG-CTTCTTGAGCAATTWAAACTCTGCCAGCACGCGACTCACTCTGTTAAGAGA 1092
```

```
Qy   301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTTCTACAGCAGCGTTCTC 360
             |||| ||||||||||| |||||  ||||  ||||||||||||||||||||||||||||||
Db   448 AACTTCCTATGCAAGGCAGTCCATGTCATCTACACAGTCAACCTTCTACAGCAGCGTCCTC 507

Qy   361 ATCCTGGCCTTCATCAGCTGACCGGTACCTGCCCATTGTCACGCCACCAACAGTCAA 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 ATCCTGGCCTTCATCAGCTGACCGGTACCTGCCCATTGTCACGCCACCAACAGTCAG 567

Qy   421 AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGCGTCTGGATCCCGCCCTC 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 AGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTTGGCCGTCTGGATCCCGCCCTC 627

Qy   481 CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540
             |||||||||||||||||||||||||||||||||||||||            ||||||
Db   628 CTGCTGACTATTCCCGACTTCATCTTTGCCAACG-----------TCAGTGAGGCA 672

Qy   541 GATGACAGGTACATCTGTGACCGGCCTTACCCGTGATAGGCTGTGTGGATGGTGGTTTCAA 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   673 GATGACAGATATATCTGTGACCGGCTTCTACCCCAATGACTTGTGGGTGGTTGTTGTTCAG 732

Qy   601 TTCCAGCATATAATGGTGGGTCTGTCATCCTGCCCGGCGATCCGTCATCCTCCTCCTGTTACTGC 660
             ||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   733 TTTCAGCACATCATGGTTGCCTTATCCTGTGTATTGTCATCCTGTCCTGCTATTGC 792

Qy   661 ATCATCATCCTCTAAGCTGTCACACTCCAAGGGCCACCAGGAAGCGCAAGGCCCCTCAAGACG 720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   793 ATTATCATCCTTCCAAGCTGTCACACTCCAAGGCCACCAGGAAGCGCAAGGCCCCTCAAGACC 852

Qy   721 ACAGTCATCCTCATCCTAGCTTCTTGCCTGCTGGCTGCCATTATTATGTGGGATCAGC 780
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   853 ACAGTCATCATCCTGGCTTTCTGCCTGTTGGCCTGTTGCCTTACTACATTGGGATCAGC 912

Qy   781 ATCGACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG 840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   913 ATCGACTCCTTCATCCTCCTCGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 972

Qy   841 CACAAGTGGATCTCCATCACACAGAGGCCCTGCCTTCTTCCCACTGTTGCCTGAACCCCATC 900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   973 CACAAGTGGAT-TTCATCACCGAGGCC--TAGCTTCTTCTTCACTGTTGTCTGTGAACCCCATCT 1029

Qy   901 CTCT 904
             ||||
Db  1030 CTAT 1033
```

RESULT 10
BX418530
LOCUS   BX418530   1201 bp   mRNA   linear   EST 13-MAY-2003
DEFINITION   BX418530 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
               CS0DF006YO16 5-PRIME, mRNA sequence.
ACCESSION   BX418530
VERSION   BX418530.1 GI:30658806
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
               Genoscope - Centre National de Sequencage
               BP 191 91006 EVRY cedex - France
               Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
               Library was constructed by Life Technologies, a division of
               Invitrogen. This sequence belongs to sequence cluster 1272.r For
               more information about this cluster, see
               http://www.genoscope.cns.fr/
               cgi-bin/cluster.cgi?seq=CS0DF006DE08QPi&cluster=1272.r. Contact :
               Feng Liang Email : fliang@lifetech.com URL :
               http://fulllength.invitrogen.com/ InVitroGen Corporation 1600

```
Qy   961 AGCAGAGGCTCCAGCCTC 978
             ::|||
Db  1093 GGGTCMRSCTCAGATCTC 1110
```

RESULT 9
BX462627
LOCUS   BX462627   1201 bp   mRNA   linear   EST 22-MAY-2003
DEFINITION   BX462627 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
               clone CS0DH005YK10 5-PRIME, mRNA sequence.
ACCESSION   BX462627
VERSION   BX462627.1 GI:31029445
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
               Genoscope - Centre National de Sequencage
               BP 191 91006 EVRY cedex - France
               Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
               Library was constructed by Life Technologies, a division of
               Invitrogen. This sequence belongs to sequence cluster 1272.r For
               more information about this cluster, see
               http://www.genoscope.cns.fr/
               cgi-bin/cluster.cgi?seq=CS0DH005BF05QP1&cluster=1272.r. Contact :
               Feng Liang Email : fliang@lifetech.com URL :
               http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
               Faraday Avenue Genoscope sequence ID : CS0DH005BF05QP1.
FEATURES   Location/Qualifiers
     source   1..1201
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="CS0DH005YK10"
               /tissue_type="T CELLS (JURKAT CELL LINE)"
               /cell_line="JURKAT CELL LINE"
               /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
               /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
               with a NotI-oligo(dT) primer. Five prime end enriched,
               double-strand cDNA was digested with Not I and cloned into
               the Not I and EcoRV sites of the pCMVSPORT 6 vector.
               Library was not normalized."
ORIGIN
Query Match   38.2%; Score 647.6; DB 13; Length 1201;
Best Local Similarity   84.8%; Pred. No. 3.1e-96;
Matches 767; Conservative   0; Mismatches 119; Indels 18; Gaps   3;

Qy   1 ATATACACTTCTGATAATACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAACAG 60
             |||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   148 ATATACACTTCAGATAATAACTACACCGAGGAAATGGGCTACATAGCTATGACTCTGAAG 207

Qy   61 GAACCCTGCTTCCCGGATGAAAACGTCCATTCATTCAATGGATTGGTGATCCTGGTCATGGGTTAC 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 GAACCCGTTTCCGTGAAGAAAAATGCTAATTTCAATAAAAATCTTCCTGCCACCATCTAC 267

Qy   121 TTCATCACTCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC 180
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 TCCATCATCTTCTTAACTGGCATCTGGGGCCAATGGATTGGTCATCCTGGTCATGGGTTAC 327

Qy   181 CAGAAGAAGCTAAGGAGCATGAGCGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 CAGAAGAAACTGAGAAGCATGAGCGGACAAGTACAGCAGCTGCACCTCCAGTCAGTGGCCGACCTC 387

Qy   241 CTCTTTGTCATCACACTCCCCTTCTGGGACCAGTTGATGCCATGGCTGCTGACTGGTACTTTGGG 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 CTCTTTGTCATCACGCTTCCCTTCTGGGCCAGTTGGTCGATGCCGTGGGCAAACTGTGTACTTTGGG 447
```

```
Qy    780  CATGCACTCCTTCATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGT  839
           |||||||||||||||||||||||| || :|||||||||||||||||||||||||||||||
Db    923  CATGCGACTCCTTCATCCTCCT-GGAAAYATCAAGCAAGGGTGTGAGTTGAGAACACTGT  981

Qy    840  GCACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTTGCCTGACCCCAT   899
           |||||||||||||||||||||||||||  :|: |||:|| ||||:|:||||||||||||
Db    982  GCACAAGTGAT-TCCATCACCGAGGCC--TAGCTTCYCMCTGTTGTCTGACCCATC    1038

Qy    900  CCTCT  904
           |||||
Db   1039  TCTAT  1043
```

RESULT 11
AA182270
LOCUS   AA182270   793 bp   mRNA   linear   EST 06-JAN-1997
DEFINITION   mt83f01.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
           IMAGE:636505 5' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
           RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION   AA182270
VERSION   AA182270.1   GI:1765810
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 793)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
           Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
           Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
           Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
           Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
           WashU-HHMI Mouse EST Project
           Washington University School of Medicine
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
           Tel: 314 286 1800
           Fax: 314 286 1810
           Email: mouseest@watson.wustl.edu
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           MGI:388497
           Seq primer: -28M13 rev2 from Amersham
           High quality sequence stop: 497.
FEATURES   Location/Qualifiers
source   1..793
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="C57BL/6J"
           /db_xref="taxon:10090"
           /clone="IMAGE:636505"
           /sex="male"
           /tissue_type="lymph node"
           /dev_stage="4 weeks"
           /lab_host="DH10B"
           /clone_lib="Soares mouse lymph node NbMLN"
           /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)
           with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
           1st strand cDNA was primed with a Not I - oligo(dT) primer
           [5'
           TGTTACCAATCTGAAGTGGGAGCGGCGGCCGCGATACTTTTTTTTTTTTTTTTTT
           3'], double-stranded cDNA was ligated to Eco RI adaptors
           (Pharmacia), digested with Not I and cloned into the Not I
           and Eco RI sites of the modified pT7T3 vector. RNA
           provided by Dr. Bertrand Jordan. Library constructed and
           normalized by Bento Soares and M.Fatima Bonaldo."
ORIGIN

Query Match   36.7%;   Score 621.6;   DB 9;   Length 793;
Best Local Similarity   92.4%;   Pred. No. 6.1e-92;
Matches 715;   Conservative 0;   Mismatches 46;   Indels 13;   Gaps 6;
```

---

```
Faraday Avenue Genoscope sequence ID : CSODF006DE08QP1.
FEATURES   Location/Qualifiers
source     1..1201
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="CS0DF006YJ16"
           /tissue_type="FETAL BRAIN"
           /dev_stage="fetal"
           /clone_lib="Homo sapiens FETAL BRAIN"
           /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
           was primed with a NotI-oligo(dT) primer. Five prime end
           enriched, double-strand cDNA was digested with Not I and
           cloned into the Not I and EcoRV sites of the pCMVSPORT 6
           vector. Library was not normalized."
ORIGIN

Query Match       36.9%;   Score 624.6;   DB 13;   Length 1201;
Best Local Similarity   84.3%;   Pred. No. 1.8e-92;
Matches 763;   Conservative 4;   Mismatches 118;   Indels 20;   Gaps 5;
```

```
Qy     1  ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG   60
          |||||| |||||| |||||||  ||||||||||| |||||||||||||||||| ||||||
Db   158  ATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGAYTCCATGAAG  217

Qy    61  GAACCCTGCTTCCCGGATGAAAACGTCCATTTCAATTCAGGGGAAAATGGGCTCAGGGGA  120
          |||||||||| ||||||||||||||| ||||||||||||||||| |||||||||||||||
Db   218  GAACCCCGTTTCCGTGAGAGAAAATGCTAATTCAATTAAAATCTTCCTGCCCACCATCTAC  277

Qy   121  TTCATCATCTTCTTGACTGCGATAGTCGGCAATGGATTGGTGATCGTCGGTCATGGGTTAC  180
          ||||||| ||||  | |||||||||  |||||||||| ||||||||| ||||||||||||
Db   278  TCCATCATCTCTTAACTGGCATTGTGGGCAATGGATTGGTCATCCTGGTTCATGGGTTAC  337

Qy   181  CAGAAGAAGTCAAGGAGCAATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC  240
          |||||||||  |||||||||||||||||||||  |||||| ||||||| ||||||||||
Db   338  CAGAAGAAACTGAGAAGCAATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCCGACCTC  397

Qy   241  CTCTTTGTCATCACACTCCCCTTCTGGCAGTTGATGCCATGGGCTGACTGGTACTTTGGG  300
          ||||||||||| |||||| ||||| |||||||||||||| ||||||| ||||||||||||
Db   398  CTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCGGTGGCAAACTGGTACTTTGGG  457

Qy   301  AAATTTTGTGTAAGGCTGTCCATATCATCTGTCAACCTCTACAGCAGCGGCGTTCTC    360
          ||||| |||||||||||| || ||||||||| ||| |||||||||||||||||| ||
Db   458  AACTTCCTATGCAAGGCAGTCCATGTCATCTCACACAGTCACCTCTACAGCAGTGTCCTC  517

Qy   361  ATCCTGGCCTTCATCAGCCTGACCAGGCTACCGGTACCTCGGCCATTGTCCACGCACCAA  420
          |||||||||||||||||  ||||||||||  |||||||||| || |||| ||||||||||
Db   518  ATCCTGGCCTTCATCAGTCTGACCAGGCTACCGGTACCTCGGCCACCAACAGTCAG  577

Qy   421  AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTTATGTGGGCCTCGTGGATCCCAGCCCTC  480
          ||||||||||||||||||||||||||| || |||||||||| |||||| |||||||||||
Db   578  AGGCCAAGGAGAAGCTGTTGGCTGAAAAGGTGGTCTTATGTTGGCGTCGGATCCCTGCCCTC  637

Qy   481  CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCGGGGGGGACATCAGTCAGGGG  540
          ||||||||| |||| ||||||| ||||||||||||| |||
Db   638  CTGCTGACTATTCCCGACTTCATCTTTCTTGCCAACG--------TCAGTRAGGCA  682

Qy   541  GATGACAGGTACATCTGTGACCGCCCTTACCCCGATAGCTGTGGATGGTGGTGTTTCAA  600
          |||||||||  |||||||||| ||||||| |||| |||||| |||| ||||||||||||
Db   683  GATGACAGATATATCTGTGACCGCCGTTCACCCCCAATGACTTGTGGGTGGTGTGTTCCAG  742

Qy   601  TTCCAGCATATAATGGTGGGGTCTCATCCTGCCGGCATCGTCCTGTCTCCTGTACTGC  660
          |||||||| ||| ||||  |||||||||||| | ||||||||||||| |||| ||||
Db   743  TTTCAGCACATCATCATGGTTGGCCTTATCCTGCTTGGCCTTATTGTCATCCTGTCCTGCTATTGC  802

Qy   661  ATCATCATCTCTAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACG  720
          |||||||||| |||| ||| |||||| |||||||| ||||| |||||||||||| |||||
Db   803  ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGACC  862

Qy   721  ACAGTCATCATCCTCATCCT-AGCTTCTTCGCCTCGGCTCGGCCTGCCATATTATGTGGGGATCAG  779
          ||||||||| |||||||||  ||||||||||| | ||||| |||||| |||  |||||||||||
Db   863  ACAGTCATCCTCATCCTCGGGCTCGGCCTGTTGGCTCGGCCTTACTACATTGGGGATCAG  922
```

TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
             WashU-HHMI Mouse EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:609995
             Seq primer: -28m13 rev2 ET from Amersham
             High quality sequence stop: 510.
FEATURES   Location/Qualifiers
     source   1..639
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C3H"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1120659"
                     /cell_line="C2C12"
                     /lab_host="DH10B"
                     /clone_lib="Barstead mouse myotubes MPLRB5"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCGGATGGAGCGGCCGCCTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [AATTCGGATCCTG], digested with Not I and cloned into the
                     Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead. The C2C12 cell line
                     (available from ATCC, catalog # CRL-1772) differentiates
                     rapidly, forming contractile myotubes and producing
                     characteristic muscle proteins."
ORIGIN

Query Match   36.7%; Score 621.2; DB 9; Length 639;
Best Local Similarity   99.5%; Pred. No. 7.4e-92;
Matches 623; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   548 GGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAATTCCAGC 607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   14 GGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTTCAATTCCAGC 73

Qy   608 ATATAATGGTGGGTCTCATCCTCGCCGGCATCGTCATCCTCCTCCGTTACTGCATCATCA 667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74 ATATAATGGTGGGTCTCATCCTCGCCGGCATCGTCATCCTCCTCCGTTACTGCATCATCA 133

Qy   668 TCTCTAAGCTGTCACACTCCAAGAGGCCACCAGAGGCCCAAGGCCCTCAAGACGACAGTCA 727
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 TCTCTAAGCTGTCACACTCCAAGAGGCCACCAGAGGCCCAAGGCCCTCAAGACGACAGTCA 193

Qy   728 TCCTCATCCAGCTTTCTTTGCCTGCTGCCCATATTATGTGGGGATCAGCATCGCGACT 787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194 TCCTCATCCAGCTTTCTTTGCCTGCTGCCCATATTATGTGGGGATCAGCATCGCGACT 253

Qy   788 CCTTCATCATCCTTTGGGAGTCATCAAGGATGCATCAAGGATGCATTGTGCACAAGT 847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 CCTTCATCATCCTTTGGGAGTCATCAAGGATGCATCAAGGATGCATTGTGCACAAGT 313

Qy   848 GGATCCATCCATCACAGAGGCCCTCGCCTTCTTCCCACTGTTGCCTGAACCCCATCCTCTATG 907
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314 GGATCCATCCATCACAGAGGCCCTCGCCTTCTTCCCACTGTTGCCTGAACCCCATCCTCTATG 373

Qy   908 CCTTCCTGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCCATGCATCCATGAGCAGCAGAG 967
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   374 CCTTCCTGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCCATGCATCCATGAGCAGCAGAG 433

Qy   968 GCTCCAGCTCGAAGATCCTTTCCAAAGGAAGAAGCGGGGGTGGACACTCTTCCGTCCCACGG 1027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   434 GCTCCAGCTCGAAGATCCTTTCCAAAGGAAGAAGCGGGGGTGGACACTCTTCCGTCCCACGG 493

Qy   493 CCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTCAGTGGGGGATGACAGTAC 552
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 CCTGACTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTCAGTGGGGGATGACAGTAC 60

Qy   553 ATCTGTGACCGCCTTTACCCCGATAGCCTGTGATGGTGGTGTTTCAATTCCAGCATATA 612
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATCTGTGACCGCCTTTACCCCGATAGCCTGTGATGGTGGTGTTTCAATTCCAGCATATA 120

Qy   613 ATGTGGGTCTCATCCTCTCCTGTTACTGCATCATCATCTCT 672
          ||||||||||||||||||||||||||||||||||||||||
Db   121 ATGTGGGTCTCATCCTCTCCTGTTACTGCATCATCATCTCT 180

Qy   673 AAGCTGTCACACTCCAGAGGCGCCACCAGAGCCGCAAGGACGACAGTCATCCTC 732
          |||||||||||||||||||| ||||||||||||||||||||||||| ||||||
Db   181 AAGCTGTCACACTCCAAGAGGCCACCAGAGCCGCAAGGACGACAGT-ATCCTC 239

Qy   733 ATCCTAGCTTCTTTGCCTGCTGCCATATTATGGGGATCAGCATCGCGACTCCTTC 792
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   240 ATCCTAGCTTCTTTGCCTGCTGCCATATTATGGGGATCAGCATCGCGACTCCTTC 299

Qy   793 ATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGACACAGTCGTGCACAAGTGGATC 852
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   300 ATCCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGACACAGTCGTGCACAAGTGGATC 359

Qy   853 TCCATCACAGAGGCCCTCGCCTTCTTCTTCCCACTGTTGCCTGAACCCCATCCTCTATGCCTTC 912
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360 TCCATCACAGAGGCCCTCGCCTTCTTCTTCCCACTGTTGCCTGAACCCCATCCTCTATGCCTTC 419

Qy   913 CTGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCATCCAACTCCATGAGCAGAGGCTCC 972
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   420 CTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCATCCAACTCCATGAGCAGAGGCTCC 479

Qy   973 AGCCTCAAGATCCTTCCAAAGGAAAGGCGGGGTGGACACTCTTCCGTCCCACGGAGTCA 1032
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   480 AG-CTCAAGATCCTT-CAAAGGAAAGGCGGGGTGGACACTCTTCCGTCCCACGGAGTCA 537

Qy   1033 GAATCCTCCAGTTTCACTCCAGCTAACCCTATGCAAAGACTTATATA---ATATATA 1088
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   538 GAATCCTCCAGTTTCACTCCAGCTAACCCTATGCAAAGACATATATATAATATATATA 597

Qy   1089 TATATAATGATAAAGAACTTTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCA 1148
          |||||||||||||||||||||||| ||||||||||||||||||||| |||||||||
Db   598 TATATAATGATAAAGAAGCTTTTTATGTTACACATTTTCCAGATATANTAGAGACTGACCC 657

Qy   1149 GTCTTGTACAGTTTTTTTTTTTTTTTTTAATTGACTGTTGTGGAGTTTATGTTTCCTCTCAGTTT 1208
          |||||||||||| |||| |||||||||| ||||||||||||||||| |||||| ||
Db   658 AGTCTGTACAGTTNNTGTTNNT----AATGACGTGGTGGGAGTTNATG-TCCTCTANGTT 711

Qy   1209 TTGTGAGGTTTGACTTAATTTATTATATAAATATTGTTTTTTGTTCATGTGA 1262
          ||||||||||||||||||| |||||||||||||||||| |||| ||||
Db   712 TTGTGAGGTTTGACTAAGTTTATTATATAAATACCTGTNNTTGTTGNTNGNTCANGTGA 765

RESULT 12
AA816049
LOCUS   AA816049   639 bp   mRNA   linear   EST 13-FEB-1998
DEFINITION vr14h02.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
            IMAGE:1120659 5' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
            RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION   AA816049
VERSION   AA816049.1 GI:2885545
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 639)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.

```
Qy  181  CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCGTCAGTGGCTGACCTC  240
              |||||||| || ||||||||||||||  |||||||||||||||||||||||||||
Db  321  CAGAAGAAACTGAGAAGCATGACGGACAAGTACCGGACGCTGCACCGTCAGTGCGGNCCTC  380

Qy  241  CTCTTTGTCATCACACTCCCCTTCTGGGCAGTCAGTTGATGCCATGGTACTTTGGG  300
              |||||||||||||     ||||||||||||||||||||||||||||||||||
Db  381  CTCTTTGTCATCACGGCTTCCCCTTCTGGGCAGTCAGTTGATGCCATGGTACTTTGGG  440

Qy  301  AAATTTTGTGTAAGGTGTCATATCATCATCATCTCACAGCAGCCGGTCTC  360
              |||||||||||  |||||||||||||||||||||||||||||||||
Db  441  AACTTCCTATGCAAGGCAGTCCATGTCATCATCACAATCAACCTTCACAGCAGTGTCCTC  500

Qy  361  ATCCTGGCCTTCATCAGCCTGGACCCGGTACCTCGCCATTGTCCACGCCACCACAGTCAA  420
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  501  ATCCTGGCCTTCATCAGTCTGGACCCGGTACCTCGCCATCGTCCACGCCACCACAGTCAG  560

Qy  421  AGGCCAAGGAAACTGCTGGCTGAAAAGGCAGTCTATGTGGGGGTCTGGATCCCAGCCCTC  480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  561  AGGCCAAGGAAGCTGTTGGCTGAAAAGGKGKGKCTATCTTGGCGTCTCGGATCCCTGCCCTC  620

Qy  481  CTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG  540
              |||||||||||||||||||||||||||
Db  621  CTGCTGACTATTCCCGACTTCATCTTTGCCAACG--------TCAGTAAGGCA  665

Qy  541  GATGACGGTACATCTGTGACCGCCCTTACCCCGATAGCCTGTGGATGGTGGTGTTCAA  600
              |||||||||||| ||||||||  ||||||||||||||||||:|||||  |||||||
Db  666  GATGACAGATATATCTGTGAMCGCTTCACCCCAATGACTTGTGGGKGKGKGXTGTGTTCCAG  725

Qy  601  TTCCAGCATATAATGGTGGGCTCATCCTGCCCGGCATCGTCATCCTGGTATTGTCATCCTGT CCGCTATTGC  660
              |  :|||||||||||||||||||||||||||||||||||| |||||||:||||||
Db  726  TTTMAGCAMATMATGGTTGGCCTATCCTGCCTGGTATTGTCATCCTGTCCGCTATTGC  785

Qy  661  ATCATCATCTTCTAAGCTGTCACACTCCAAGGGCCAAGGGCCAAGCGCAAGACG  720
              |||||||||   |||||||||||||   ||||||||||||||||||||||
Db  786  ATTATCATCTCCAAGCTGTCACACTCCAAGGGSCACCAGAAGCGCAAGGCCCTCAAGACC  845

Qy  721  ACAGTCATCCTATCCTATCCTAGCTTTCTTTGCCTGCTCGCTGTCGCCGATCAGC  780
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db  846  ACAGTCATCATCCTCATCCTCGGCTTTCTTGCCTGCTCGCCTTACTACATTGGATMA-S  904

Qy  781  ATCGACTCCTCATCCTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTGTG  840
              || |||||||||||||||||||  ||||| | ||||:||||:|||||||||
Db  905  ATCGACTCCTTCATCCTCCTGGAAATCATMAAGCAAGGGKGKGTGAGTTTRARA-MACTGTK  963

Qy  841  CACAAGTGGATCTCCATCACAGAGAGGCCCTGGCCTTCTTCCCACTGTTGCCTGAACCCCATC  900
              |||:|:| ||||||:| ||||:|||:||:| |||:|:|: ||:|:||||||
Db  964  CAMAATGRTT-----TCAYAMCGAGGCYARCTTYTTYTCMWTKTYKTTTKACCCAHCYC  1018

Qy  901  CTCTATGCCTT  911
              ||||||||
Db  1019  TWTGTTTYTTT  1029
```

RESULT 14

BX362477

LOCUS	BX362477 952 bp mRNA linear EST 05-MAY-2003
DEFINITION	BX362477 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED Homo sapiens cDNA clone CS0DJ015YP09 5-PRIME, mRNA sequence.
ACCESSION	BX362477
VERSION	BX362477.1 GI:30380640
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 952)
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and normalization
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope
	Genoscope - Centre National de Sequencage
	BP 191 91006 EVRY cedex - France
	Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

```
Qy  1028  AGTCAGAGAATCCTCCAGTTTCACTCCAGTCAACCCTTATGCAAGAGACTTATATAATATAT  1087
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  494  AGTCAGAGAATCCTCCAGTTTCACTCCAGTCAACCCTTATGCAAGAGACTTATATAATATAT  553

Qy  1088  ATATATATATGATAAAGAACTTTTTATGTTACACATTTCAGATATAAGAGACTGACC  1147
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  554  ATATATATATGATAAAGAACTTTTTATGTTACACATTTCCAGATATAAGAGACTGACC  613

Qy  1148  AGTCTTGTACAGTTTTTTTTTTTTT  1173
               ||||||||||||||||||
Db  614  AGTCTTGTACAGTTTTTTTTTTTTTT  639
```

RESULT 13

BX462182

LOCUS	BX462182 1201 bp mRNA linear EST 22-MAY-2003
DEFINITION	BX462182 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA clone CS0DG003Y019 5-PRIME, mRNA sequence.
ACCESSION	BX462182
VERSION	BX462182.1 GI:31031389
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1201)
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and normalization
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope
	Genoscope - Centre National de Sequencage
	BP 191 91006 EVRY cedex - France
	Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
	Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1272.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DG003AH10QP1&cluster=1272.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Genoscope sequence ID : CS0DG003AH10QP1.
FEATURES	Location/Qualifiers
source	1..1201
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CS0DG003Y019"
	/tissue_type="B CELLS (RAMOS CELL LINE)"
	/cell_line="RAMOS CELL LINE"
	/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
	/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match   35.9%;   Score 608.2;   DB 13;   Length 1201;
Best Local Similarity   80.1%;
Matches 730;   Conservative 29;   Mismatches 130;   Indels 22;   Gaps 4;

```
Qy  1    ATATACACTTCTGATAACTACTCTGAAGAAGTGGGCTCTGGAGACTATGACTCCAACAAG  60
              |||||||||||||||||||||||||||||||| |||||||||||||||
Db  141  ATATACACTTCAGATAACTACACACCGAGGAAATGGGCTCAGGGGGACTATGACTCCATGGAAG  200

Qy  61   GAACCCTGCTTCCCGGATGAAAACGTCCATTTCAATAGGATCTTCCTGCCCACCATCTAC  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201  GAACCCTGTTTCCGTGAAGAAAATGCTAATTCAATTCATTAAAAATCTTCCTGCCCACCATCTAC  260

Qy  121  TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCATGGGTTAC  180
              ||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  261  TCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTGATCCTGGTCATGGGTTAC  320
```

```
Db    799 ATTATCATTCCAAGCTGTCACACTCCAAGAGGGCCACCAGGCCAAGCCCTCAAGACC 858
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    721 ACAGTCATCCTCATCCTAGCTTTCTTTGCCTGCTGGCTGCTGCCATTATGTGGGATCAGC 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    859 ACAGTCATCCTCATCCTGGCTTTCTTCTCGGCCTGTTGGCTGCCTTACTACATTGGGATCAGC 918

Qy    781 ATGGACTCCTTCATCCTTTTGGGA 804
           ||||||||||||||| :.|:.|:
Db    919 ATCGGACTCCTTCATCYYCTKGSAA 942
```

RESULT 15
CD251331
LOCUS   CD251331   872 bp   mRNA   linear   EST   22-MAY-2003
DEFINITION   AGENCOURT_14211919 NIH_MGC_179 Homo sapiens cDNA clone
             IMAGE:3038274 5', mRNA sequence.
ACCESSION   CD251331
VERSION   CD251331.1   GI:31011797
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 872)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
             Office of Cancer Genomics
             National Cancer Institute / NIH
             Bldg. 31 Rm10A07 Bethesda, MD 20892
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: NDAM449 row: h   column: 11
             High quality sequence stop: 654.
FEATURES   Location/Qualifiers
     source   1..872
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3038274"
                     /tissue_type="Pituitary"
                     /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_179"
                     /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
                     (destroyed); Site_2: NotI; Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.1 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."

ORIGIN

Query Match   35.5%;   Score 602.2; DB 14; Length 872;
Best Local Similarity 85.9%;   Pred. No. 8.9e-89;
Matches 697;   Conservative   0;   Mismatches   98;   Indels 16;   Gaps   2;

```
Qy    159 GGTGATCCTGGTCATGGGTTACCAGAGAAGCTAAGGAGCATGACCGGACAAGTACCGGCT 218
           |||  |||||||||||||| |||| ||||||||||||||   |||||||||||||||||||
Db      1 GGTCATCCTGGTCATGGTTACCAGAGAAGCTAAGGAGAGAAACTGAGAGCATGACGGACAAGCAGGCT 60

Qy    219 GCACCTGTCAGTGGCTGACCCCTTCTTTGTCATCACCACTCCCCTTCTGGGCAGTTGATGC 278
           |||||||||||||||||||||||||||  ||||||||||||||| |||||||||||||||||
Db     61 GCACCTGTCAGTGGCTGGCCGACCTCTTTGTCATCACCACGCTTCCCTTCTGGGCAGTTGATGC 120

Qy    279 CATGGCTGACTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCCTACACTGT 338
           |||| |||| || ||||||||||||| |     |||||||   |||||| ||||||||
```

```
Db   121 CGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACAGT 180

Qy   339 CAACCTCTACAGCAGCGGTTCTCATCCTGGCCTTCATCGACCTGGACCGGTACCTCGCCAT 398
         ||||||||||||| ||||||||||||||| |||||||||||||| ||||| |||| ||||
Db   181 CAACCTCTACAGCAGTGTCCTCATCCTGGCCTTCATCAGTCTGGACCGGCTACCTCGGCCAT 240

Qy   399 TGTCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGCTGAAAGGCAGTCTATGT 458
         |||||||||| |||||||||||||||||| ||| |||||||||||||||||||||
Db   241 CGTCCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGCTGAAAAAGTGGTCTATGT 300

Qy   459 GGGCGTCTGATCCCAGCCCTCCTGACTATACCTGACTTCATCTTTGCCGACGTCAG 518
         |||||||||||||| ||||||| ||||||||||||||||||||||||||
Db   301 TGGCGTCTGGATCCCTGCCCCTCCTGCTGACTATTCCCGACTTCATCTTTGCCAACG---- 356

Qy   519 CCAGGGGGACATCAGTCAGGGGATGACAGCAGTACATCTGTGACCCGCCTTACCCCGATAG 578
         |             |||||||||||||||||||| ||||||||||||||||||
Db   357 ------------TCAGTGAGGCAGATGACAGATATCTGTGACCGCTTCTACCCCAATGA 405

Qy   579 CCTGTGGATGGTGGTGTTCAATTCCAGCATATAATGGTGGGTCTCATCCTGCCCGGCAT 638
         ||||| |||||||||||| ||||||| || ||||||||||| ||||| |||||
Db   406 CTTGTGGGTGGTTGTGTTCCAGTTTCAGCACACATGGTTGGCCTTATCCTGCCTGGGTAT 465

Qy   639 CGTCATCCTCTCCTGTTACTGCATCATCATCCTAAGCTGTCACACTCCAAGGGCCACCA 698
         ||||||||| |||| | ||||||||||||||||||||||||||||||||||||
Db   466 TGTCATCCTGTCCTGCTATTGCATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCA 525

Qy   699 GAAGCGCAAGGCCCTCAAGACGACAGTCATCCTCCTAGCTTTCTTTGCCTGCTGGCT 758
         |||||||||||||| || ||||||||||| ||||||| ||||||| |||||||||
Db   526 GAAGCGCAAGGCCCTCAAGACCACAGTCATCCTCTCATCCTTCTTCGCCTGTTGGCT 585

Qy   759 GCCATATTATGTGGGGATCAGCATGCATGGACTCCTTTTGGGAGTCATCAGCAAGG 818
         |||| |||| |||||||||| ||||||||||| |||||| ||||| ||||||||||
Db   586 GCCTTACTACATTGGGATCAGCATGCATGCTTCCTTCATCCTCCTGGAAATCATCAAGCAAGG 645

Qy   819 ATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTCCATCACACAGAGAGGCCCTCGCCTTCTT 878
         ||||||| |||||||| |||||||||||| ||||| |||||||| ||||||||||||||||||
Db   646 GTGTGAGTTTGAGAACACTGTGCACAAGTGATTCCATCACCGAGGCCCTAGCTTTCTT 705

Qy   879 CCACTGTTGCCTGAACCCCATCCTCCTATGCCTTCCCTCGGG-GCCAAGTTCAAAAGCTCTG 937
         ||||||| ||||||||||| ||||| ||||||| |||||| |||||| ||||| |||||
Db   706 CCACTGTTGTCTGAACCCCATCCTCCTATGCTTTCCTTGGGAGCAAATTTTAAAACCTCTG 765

Qy   938 CCCAGCATGCACTCAACTCCATGAGCAGAGG 968
         |||||| |||| || ||||||||||||||||
Db   766 CCCAGCGACGCACTCACCTCTGTGAGCAGAGG 796
```

Search completed: May 17, 2004, 16:09:12
Job time : 4894.52 secs